10 Å cryo-EM structure of *Drosophila melanogaster* hexamerin and *Limulus polyphemus* hemocyanin

10 Å kryo-EM-Struktur von *Drosophila melanogaster* Hexamerin und *Limulus polyphemus* Hämocyanin

Dissertation
zur Erlangung des Grades

„Doktor der Naturwissenschaften“

am Fachbereich Biologie der

Andreas G. Martin
geb. am 16.09.1975 in São Paulo

Mainz, im November 2006

Tag der mündlichen Prüfung: 21. Dezember 2006

Los ordenadores son inútiles. Sólo pueden darte respuestas.

Computers are useless. They can only give you answers.

Pablo Picasso (1881 - 1973)
Spanish cubist painter

Contents

List of Figures

List of Tables

Abbreviations

1D	one-dimensional
2D	two-dimensional
3D	threedimensional
Å	Ångstrom (10^{-10} meter)
CTF	Contrast transfer function
Da	Dalton
dpi	dots per inch
EM	Electron microscop, microscopy
FSC	Fourier shell correlation
GHz	Gigahertz
HAC	Hierachical ascendant classification
`hed`	header – `IMAGIC 5`-header file
`img`	image – `IMAGIC 5`-pixel file
kDa	10^3 Dalton
LSP	Larval serum protein
MB	Megabyte
MHz	Megahertz
MRA	Multiple reference alignment
MSA	Multivariate statistical analysis
nm	Nanometer – 10^{-9} meter
pdb	Protein data bank
RAM	Random access memory
RT	Room temperature
SNR	Signal-to-noise ratio
TEM	Transmission electron microscop, microscopy
`tif`	tagged image format

Chapter 1

Introduction

About 200 years ago, Antoine Lavoisier (1743-1794) showed that the gas he termed 'oxygen' (acid generating) is consumed by animals while they produce carbon dioxide and heat. If the oxygen supply of our brain is suppressed, unconciousness occurs after a few seconds and irreparable damage is caused after three to four minutes. Oxygen is the final electron acceptor in the energy generating process of the oxidative phosphorylation. The constant supply of oxygen is crucial for the survival of cells.

To cope with this challenge, either structural features assure that each cell of the body is close enough to the breathing medium, as the trachea system (internal air channels) of e.g. the insects, or oxygen binding proteins enrich the body fluids surrounding the cells with this vital molecule. Evolution came up with three different types of oxygen binding proteins: hemoglobins, hemerythrins, and hemocyanins.

Hemoglobins, a wide spread type of oxygen binding proteins, are present in most vertebrates as well as in several species of the arthropods, molluscs, and annelids. Human hemoglobin is a rather small molecule with a molecular weight of about 64,000 that is tightly packed in the red blood cells, the erythrocytes. This hemoglobin increases the blood's capability of to transport oxygen from 4 ml/l (solved in water) to 200 ml/l, a 50-fold quantity. The dioxygen molecule is reversibly bound to an iron ion that is located in the center of a heme ring. The giant extracellular hemoglobins of some arthropods and molluscs are also termed 'erythrocruorins'.

A variation of hemoglobin are the 'chlorocruorins'. They are giant extracellular heme proteins found in four marine polychaete families ([Pallavicini *et al.*, 2001]). They contain an altered heme with a formyl group substituting the canonical three-vinyl group and consequently appear as greenish red ([Lamy *et al.*, 1996]).

Hemerythrins are mainly found in animals of several marine invertebrate phyla (e.g. spinculids, priapulids, and some brachipods and annelids), but is also present in the prokaryotic bacterium *Methylococcus capsulatus* ([Karlsen *et al.*, 2005]). The active site of this metalloproteins contain two iron atoms linked by carboxylate groups and a μ-oxo bridging atom ([Stenkamp, 1994]).

The hemocyanins will be presented in detail in the following section.

1.1 Hemocyanins

Hemocyanins (Hc) are the extracellular oxygen transport proteins of many molluscs and arthropods and freely dissolved in the hemolymph. The oxygen-binding site is formed by a dinuclear copper type 3 centre that is directly bound to the protein (Figure 1.1), causing the blue coloration of the hemolymph when oxygenated ([Redmond and Gamlin, 1999]). Hemocyanins show a complex allosteric behaviour during oxygen binding.

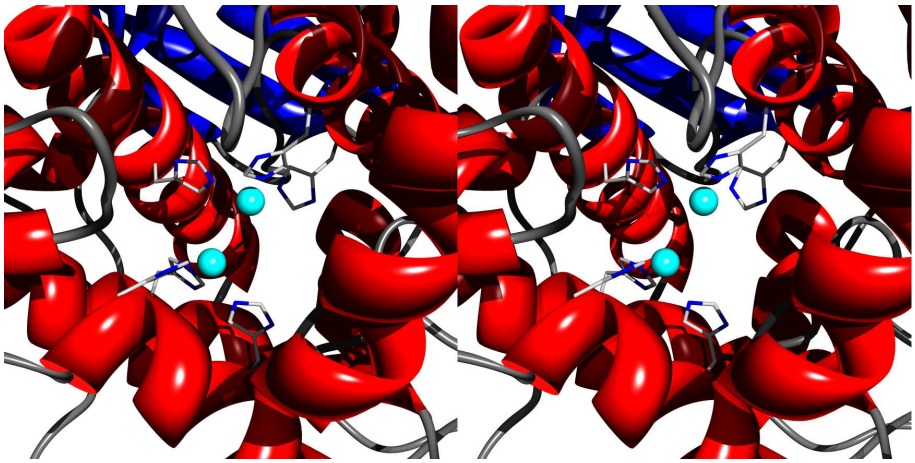


Figure 1.1: Stereo view of the dinuclear copper type 3 centre from the arthropod *Limulus polyphemus*. The two copper ions (spheres) are bound to the protein through six histidines (wireframes – PDB ID: 1NOL).

Hemocyanins form a heterogeneous group of proteins. On the level of the primary structure, the sequence similarity between molluscs and arthropod hemocyanins is limited to the region that binds the second copper ion ([Volbeda and Hol, 1989b], [Lang and van Holde, 1991]). Electron microscopic data reveals that the quaternary structures are not related in any way. The molluscan hemocyanins form huge cylindrical complexes (decamers, didecamers, or multidecamers) with M_r of up to 8,000,000 and beyond. Each cylinder consists of polypeptides of M_r 350,000 or 400,000 that carry seven or eight binuclear copper sites ([van Bruggen *et al.*, 1981], [van Holde *et al.*, 1992]).

In contrast, arthropod hemocyanins are hexamers (1×6mers) or oligohexamers (n×6mers, Figure 1.2) of polypeptides of M_r ∼75,000, each containing only a single dinuclear copper site (see review [Markl and Decker, 1992]). The polypeptides are bean-shaped and comprise ∼660 amino acids. The 'basic building block' is a hexamer with a molecular mass of ∼450,000 Da (the size range of a single polypeptide chain of the hemocyanin from molluscs) arranged as a trigonal antiprism with approximatly 32 (D3) point-group symmetry. The elucidation of the X-ray structure of spiny lobster *Panulirus interruptus* hemocyanin was a major step forward in understanding the hexameric oligomer ([van Schaick *et al.*, 1982], [Gaykema *et al.*, 1984],

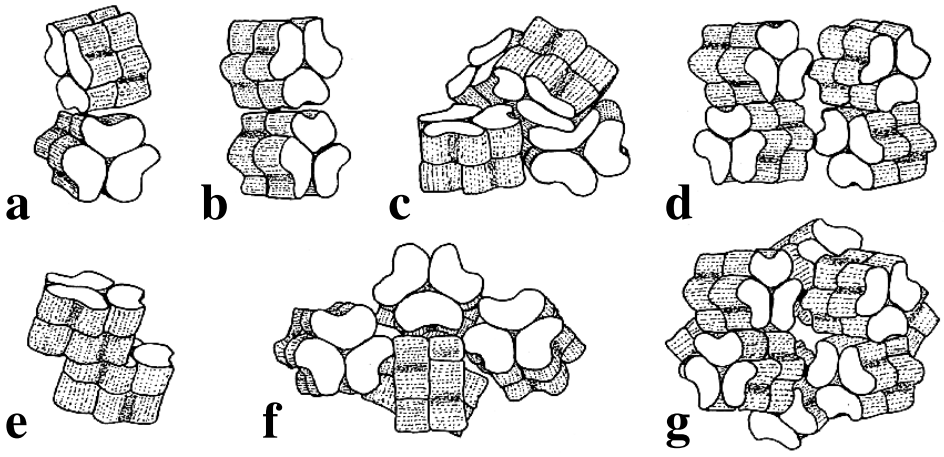


Figure 1.2: Schematic drawings of the quaternary structure of different arthropod hemocyanins according to [Markl and Decker, 1992]. **a** – typical crustacean 2×6mer, **b** – 2×6mer of certain spiders, **c** – thalassinid shrimps 4×6mer, **d** – 4×6mer of many arachnids, **e** – unusual crustacean 2×6mer found in *Squilla mantis*, **f** – 6×6mer from the centipede *Scutigera coleoprata*, **g** – 8×6mer from *Limulus polyphemus*.

[Volbeda and Hol, 1989a]). The X-ray structure of a homo-hexamer comprising only subunit type II from the horseshoe crab *Limulus polyphemus* was presented some years later and showed a very similar quaternary structure ([Magnus *et al.*, 1991], [Hazes *et al.*, 1993], [Magnus *et al.*, 1994]). Higher-ordered arthropod hemocyanins can be assembled of up to eight hexamers (8×6mers).

Single hexamers (1×6mers) are found in natantian shrimps, krill, spiny lobsters, fiddler crabs, and many isopods ([Ellerton *et al.*, 1983], [Markl, 1986]). Hemocyanin related hexamers can also be found in crustaceans ([Terwilliger *et al.*, 1999]) and insects as copper-free serum proteins (section 1.2).

The 2×6mers exist in three different variations of the quaternary structure. 2×6mers with the hexamers rotated 90° in respect to each other (Figure 1.2 a) can be found in most Reptantia, some natantian shrimps, and the isopod *Ligia* ([Terwilliger, 1982], [Ellerton *et al.*, 1983], [Markl, 1986]). A 2×6mer with two stacked hexamers overlapping to about 60% of their hexagonal surfaces (Figure 1.2 e) is present in the stomatopod *Squilla mantis* ([Bijlholt and van Bruggen, 1986]). Many labidognath spiders have 2×6mers with hexamers connected by a one-point disulfide bridge contact ([Markl, 1980], [Markl, 1986], Figure 1.2 b).

Also among the 4×6mers, variations of the quaternary structure can be observed. Hemocyanins from the thalassinid shrimps have two 2×6mers with a 90° rotation arranged as a tetrahedron (Figure 1.2 c), in which each corner is occupied by a hexamer ([van Bruggen *et al.*, 1981], [van Bruggen, 1983]). The 4×6mers of many Arachnida are built by two 2×6mers (the hexamers with a 90° rotation) in an antiparallel arrangement (Figure 1.2 d). The 2×6mers are connected via two central bridges. The centers of the hexamers are not coplanar and thus a 'rocking' effect can

be observed. Furthermore, a skewing of the dodecamers leads to a slightly rhombic shape with two different faces, known as 'flip' and 'flop' ([van Heel and Frank, 1981], [Bijlholt *et al.*, 1982]).

In the centipede *Scutigera coleoptrata*, a myriapode, a 6×6mer was found (Figure 1.2 f). A first proposal, as well as image processing and model building, suggest an approximately octahedral arrangement of three 2×6mers ([Mangum *et al.*, 1985], [Boisset *et al.*, 1990]).

The 8×6mer of the Xiphosura is the largest arthropod hemocyanin known (Figure 1.2 g). Eight different subunit types are present in a different number of copies in the 8×6mer of *Limulus polyphemus*: six copies of type I, eight copies of type II, two copies of type IIA, eight copies of type IIIA, eight copies of type IIIB, eight copies of type IV, four copies of type V, and four copies of type VI subunit ([Sullivan *et al.*, 1974], [Lamy *et al.*, 1979], [Brenowitz *et al.*, 1981], [Lamy *et al.*, 1983]). This heterogeneity correlates with the cooperativity of these oxygen transport proteins ([Markl, 1986], [Topham *et al.*, 1986], [Decker *et al.*, 1989]) and serves a prerequisite for the assembly to higher quaternary structures ([Lamy *et al.*, 1977], [Bijlholt *et al.*, 1979], [Markl, 1980], [Markl *et al.*, 1982], [Brenowitz *et al.*, 1984]).

Although a plethora of arthropod hemocyanin subunit sequences is now known ([Burmester, 2001], [Burmester, 2002]), only in a few cases the complete subunit set of a particular hemocyanin has been sequenced. Among the chelicerates, this is only the case for the tarantula *Eurypelma californicum*, the golden orb-web spider *Nephila inaurata* (with one type missing) and the ctenid spider *Cupiennius salei* ([Voit *et al.*, 2000], [Ballweber *et al.*, 2002], [Averdam *et al.*, 2003]).

Thus, many details of the primary and quaternary structure of arthropod hemocyanins are known. However, sequence data of several functionally well-studied hemocyanins are still incomplete or lacking, and the exact mode of the higher-ordered assembly of hexamers remains obscure. With respect to the 4×6mer/8×6mer chelicerate hemocyanins, different image analysis procedures were used to solve the quaternary structure of *Androctonus*, *Eurypelma* and *Limulus* hemocyanin, starting from electron microscopic images of negatively stained preparations (e.g. [van Heel and Frank, 1981], [Bijlholt *et al.*, 1982], [Lamy *et al.*, 1982], [van Heel and Dube, 1994], [de Haas and van Bruggen, 1994]).

Three-dimensional reconstruction from cryo-electron microscopical images was performed for the *Androctonus* 4×6mer ([Boisset *et al.*, 1995]) and the *Limulus* 8×6mer ([Taveau *et al.*, 1997]), but in both cases, the resolution limit was about 40 Å. These approaches revealed various basic parameters of the 4×6mer and the 8×6mer, but a reliable molecular fitting of the crystal structure was not possible due to the low resolution obtained. The exact orientation of the hexamers within the oligo-hexameric, and their exact contact regions, are still unknown. This, however, is prerequisite to fully understand the allosteric functions of these proteins.

1.2 Hexamerins

Hexamerins are hemocyanin-related storage proteins present in the hemolymph of most of the studied insects. Phylogenetic analyses have shown that the hexamerins from insects are derived from crustacean hemocyanins ([Beintema *et al.*, 1994], [Burmester and Scheller, 1996], [Burmester *et al.*, 1998]) and diversificated along with the divergence of the insect orders ([Burmester *et al.*, 1998]). They belong to the athropod hemocyanin superfamily ([Burmester, 2001]). Hexamerins have five of the six copper binding histidines replaced by other residues and thus contain no copper. The ability to reversibly bind oxygen is therefore absent. A typical hexamerin is built of six identical or closely related subunits with M_r ~80,000.

Hexamerins are tyrosine- and phenylalanin-rich serum proteins and are also referred to as arylphorins ([Telfer *et al.*, 1983]). The concentration of the storage proteins in the hemolymph can rise significantly, especially in the final instar larvae prior pupation ([Telfer and Kunkel, 1991]). It is assumed that these proteins provide energy and amino acids during metamorphosis ([Munn and Greville, 1969], [Telfer and Kunkel, 1991]). Furthermore, hexamerins are cuticle structural compounds ([Tsakas *et al.*, 1991], [Peter and Scheller, 1991], [Chrysanthis *et al.*, 1994]). They are involved in the process of humoral immunological response ([Phipps *et al.*, 1994], [Beresford *et al.*, 1997]), and in the transport of hormones ([Ismail and Gillot, 1995], [Braun and Wyatt, 1996]) and small organic compounds ([Magee *et al.*, 1994], [Miura *et al.*, 1994]).

Two immunologically distinct hexamerins have been identified in the fruitfly, *Drosophila melanogaster*: larval serum protein 1 (LSP-1) and larval serum protein 2 (LSP-2) ([Roberts *et al.*, 1977], ([Roberts, 1987], [Roberts *et al.*, 1991a], [Roberts *et al.*, 1991b]). LSP-1 is a heterohexamer built of randomly associated α-, β-, and γ-subunit types, whereas LSP-2 is a homohexamer. A comparison of the subunit types is shown in table 1.2.

LSP-1α	762 amino acids	92,9 kDa
LSP-1β	773 amino acids	94,4 kDa
LSP-1γ	756 amino acids	91,7 kDa
LSP-2	697 amino acids	83,4 kDa

Table 1.1: Comparison of the hexamerin subunit types from *Drosophila melanogaster*.

The LSP-1 and LSP-2 expression starts at the beginning of the last larval instar in the fat body. LSP-1 expression starts earlier and reaches a higher level than the expression of LSP-2 ([Powell *et al.*, 1984]), but in adults only LSP-2 is further expressed ([Benes *et al.*, 1990], [Shirras and Bownes, 1989]). LSP-1 knock-out mutants have shown that this protein is possibly dispensable for the metamorphosis, but these mutants had a lower fertility in both sexes ([Roberts *et al.*, 1991b]). Thus, LSP-1 plays either a specific role in the development of reproductive or-

gans or the exclusive expression of LSP-2 is not sufficient for supplying the needed amounts of essential amino acids for the metamorphosis. Two different types of hexamerins within *Drosophila melanogaster* suggest differnt functions, but a specific role of LSP-2 has not been identified yet ([Beverley and Wilson, 1982], [Brock and Roberts, 1983], [Beverley and Wilson, 1984], [Burmester, 1999]). The quaternary structure of hexamerins is not known in detail, although a preliminary model is available ([Markl *et al.*, 1992]).

1.3 Single particle analysis

The invention of the electron microscope (EM) by Ernst Ruska (Figure 1.3) in the early 1930ies provided a new and powerful instrument for analysis in a big variety of fields, including the determination of the structure of isolated biological macromolecules (single particles). With modern electron microscopes that take advantage of the very short electron wavelength, resolution levels up to ∼0.7 Å are theoretically possible. The 'phase' problem of the X-ray crystallography is absent in the EM because it produces images of the analyzed samples rather than diffraction patterns.

However, there is a serious limitation to the resolution in the analysis of biological assemblies that roots in the radiation sensitivity of the analyzed particles. The electron beam of the microscope interacts in different ways with the sample during exposure. The image is created by elastically scattered electrons which leave their 'prints' on a photographic emulsion. During the collection of the data, non-elastic interactions gradually damage the particles. To reduce the damaging effect of non-elastically scattered electrons, imaging is performed under low-dose and cryo conditions.

Under low-dose conditions (5–10 electrons/Å^2), the signal-to-noise ratio (SNR) of the images gets very poor and information is veiled by noise, but beam damage is reduced to tolerable levels. Freezing the sample in vitrious ice and cooling it to the temperature of liquid nitrogen (∼100 K) or even lower (to the temperature of liquid helium close to 0 K) permits an increase of the exposure dose by the factor 10–30 ([Stark *et al.*, 1996]) and thus an improvement of the SNR.

The rather poor quality of the images can be improved by averaging techniques. Therefore, a large number of individual images of macromolecular assemblies is required. The acquisition of these images can be divided in 'zero-tilt' and 'tomography' ([Hoppe *et al.*, 1974], [Dierkensen *et al.*, 1992], [Radermacher *et al.*, 1987], [Radermacher, 1988]) techniques. The tomography technique tilts the sample in the microscope producing images of the same particles in different and defined orientations. For the zero-tilt technique, the random orientation of the particles embeded in the vitrious ice is used to access the 3D information of the macromolecule. The 3D information can be extracted in two different approaches: the 'projection matching' an the 'angular reconstitution'. In projection matching, re-

Figure 1.3: Winner of the Nobel prize Ernst Ruska (right) and his assistant Max Knoll adjusting the electron microcope (source http://www.tu-berlin.de/uebertu/fotogalerie/historisch/slides/ruska.html).

projections of a starting model or an earlier cryo-EM structure are matched to the images by correlating them ([van Heel, 1984b], [Harauz and Ottensmeyer, 1984], [Harauz and van Heel, 1986a]). Angular reconstitution is based on finding common line projections between 2D projection images of the same 3D object ([van Heel, 1987], [Orlova and van Heel, 1994]).

The strength of the single particle analysis has already been demonstrated in a number of studies by reaching sub-nanometer resolutions (e.g. [Böttcher *et al.*, 1997], [Meissner *et al.*, 2003], [Ludtke *et al.*, 2004], [Klaholz *et al.*, 2004], [Rawat *et al.*, 2006]). Resolution levels of ∼4 Å, at which the backbone of the polypeptide chains would be identified allowing an interpretation at atomic level, appear to be possible in near future by this technique.

The following study is based on the zero-tilt technique with the combination of angular reconstitution (as first approach and in the initial stages of the reconstruction) and projection matching (for final refinements).

1.4 Objectives

Hexamerin LSP-2 is a hexameric hemolymph protein, constituted of six identical ∼83 kDa subunits. The quaternary structure, especially of the LSP-2 from *Drosophila melanogaster*, has long time been nebulous, because the EM-micrographs only showed cubic shaped structures ([Mousseron-Grall *et al.*, 1997]). The cubic shape was also observed in the closely related LSP-2 from *Calliphora vicina* ([Markl *et al.*, 1992]).

The question was whether and to which degree the structure of a trigonal antiprism with 32 (D3) point-group symmetry that is typical for arthropod hemocyanins is also applicable to the insect hexamerins ([Markl *et al.*, 1992]).

Hemocyanin The function of a protein is strongly correlated with its structure. Arthropod hemocyanins show a moderate to low oxygen affinity, high cooperativity, and an accentuated Bohr effect ([Loewe, 1978], [van Holde and Miller, 1982], [Ellerton *et al.*, 1983]). The hemocyanin of *Limulus polyphemus* employs cooperative oxygen binding and heterotropic allosteric regulation by protons, chloride ions and divalent cations ([Brouwer *et al.*, 1977]), and its quaternary structure is stabilized by hydrophilic and polar forces ([Dolashka-Angelova *et al.*, 2005]). To interpret the functionality of this huge oxygen transport machine, detailed knowledge of the quaternary structure is fundamental.

The second aim of this study was to achieve a highly resolved cryo-EM structure of the 8×6mer hemocyanin from *Limulus polyphemus* with a known oxygenation state. Furthermore, the amino acids of the subunits involved in the different types of inter-hexamer contacts within the 8×6mer should be precisely identified with the techniques of molecular modelling and docking. These results should allow an insight into how allosteric effects are transferred between the eight hexamers.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 *Limulus polyphemus*

Limulus polyphemus (Linnaeus, 1758) is classified as follows:

Kingdom	*Animalia* – animals
Phylum	*Arthropoda* – arthropods
Subphylum	*Chelicerata*
Class	*Merostomata* – horseshoe crabs
Order	*Xiphosura*
Suborder	*Limulina*
Family	*Limulidae*
Genus	*Limulus*
Species	*Limulus polyphemus*

Thirteen specimens of living *Limulus polyphemus* were obtained from the Marine Biolab (Woodshole – USA). The horseshoe crabs were kept in seawater at 18°C with a 12h/12h light-dark cycle and fed on fish.

2.1.2 *Drosophila melanogaster*

Drosophila melanogaster (Meigen, 1830) is classified as follows:

Kingdom	*Animalia* – animals
Phylum	*Arthropoda* – arthropods
Subphylum	*Hexapoda* – hexapods
Class	*Insecta* – insects
Subclass	*Pterygota* – winged insects
Infraclass	*Neoptera* – modern, wing-folding insects

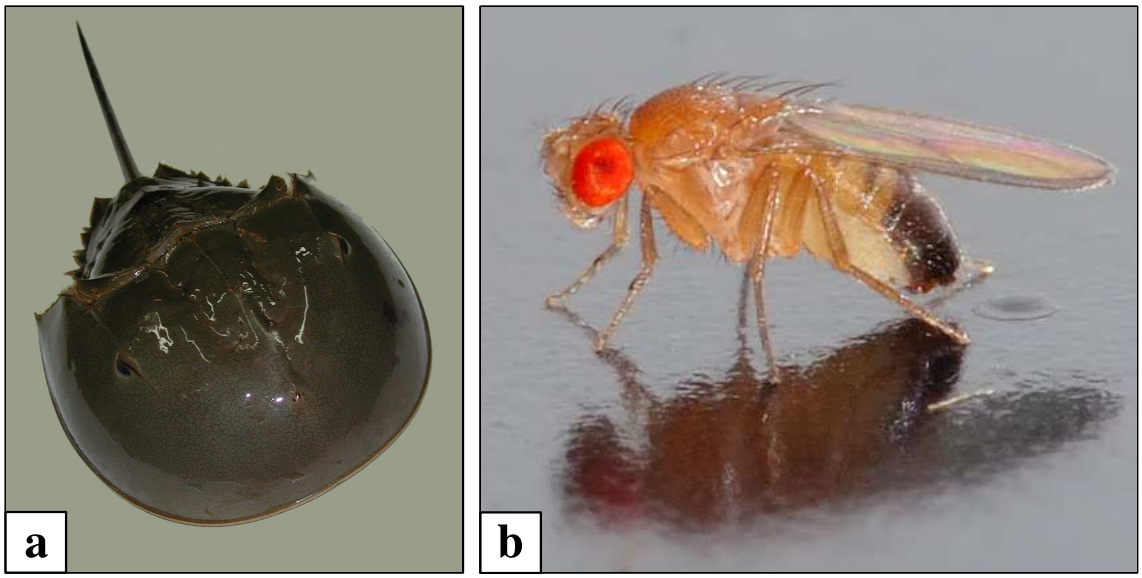


Figure 2.1: *Limulus polyphemus* (a) and *Drosophila melanogaster* (b – source http://de.wikipedia.org/wiki/Bild:Drosophila_melanogaster_-_side_%28aka%29.jpg).

Order	*Diptera* – gnats, mosquitoes
Suborder	*Brachycera* – circular-seamed flies, muscoid flies
Infraorder	*Muscomorpha*
Family	*Drosophilidae* – pomace flies, small fruit flies
Subfamily	*Drosophilinae*
Genus	*Drosophila*
Species	*Drosophila melanogaster*

The wild-type *Drosophila melanogaster* stock Oregon R was mantained at 25°C on standard yeast/cornmeal/sucrose medium sprinkled with active dry yeast. Propionic acid and Nipagi were added as inhibitors for moulds. The larvae were grown in the facilities of Institute for Genetics of the University of Mainz.

2.1.3 Chemicals and apparatuses

Analytic quality chemicals were obtained from Roth (Karlsruhe – D), Merck (Darmstadt – D) and Serva (Heidelberg – D). Apparatuses used were:

Photometer	Ultrospec Plus (Pharmacia LKB, Freiburg – D)
Centrifuge	Centrifuge 5415R (Eppendorf, Colone – D)
pH Meter	Digital-pH-Meter (Knick, Berlin – D)

Reaction vessels were obtained from Falcon (Heidelberg – D) and Eppendorf (Colone – D). Other chemicals and apparatuses are described elsewhere.

2.1.4 Antibodies

Primary antibodies	Anti-LSP-1 and anti-LSP-2, polyclonal antibodies against LSP-1 and LSP-2 hexamerins, antiserum from rabbit (kindly provided by T. Burmester)
Secondary antibodies	Goat-anti-rabbit, coupeled with basic phosphatase, IgG (H+L), obtained from Dianova (Hamburg – D)

2.1.5 Electron microscopes

Three different microscopes were used, two of them able to work under cryo conditions:

- Zeiss-EM 9A (80 keV) equipped with a wolfram cathode – Mainz – D
- Philips CM200 cryo-EM (200 keV) equipped with a field emission gun (FEG) – MPI Göttingen – D
- TECNAI F30 cryo-EM (300 keV) equipped with a FEG – Mainz – D

2.1.6 Hardware

Image processing was initially performed on a computer cluster (or 'farm') with a 32 bit `INTEL Xeon` 3.06 GHz dual processor PC (3 GB RAM) as headnode and ten 32 bit `AMD Athlon MP` 2.6 GHz dual processor PCs (512 MB RAM) as nodes. The computer cluster was acquired at Technical Point Computer GmbH (Hochheim – D)

Further processing was performed on a 'Beowulf' cluster with a 64 bit `AMD Opteron` 2.4 GHz dual core dual processor (4 GB RAM) PC as headnode and eight 64 bit `AMD Opteron` 2.2 GHz dual core dual processor (4 GB RAM) blades as nodes. The Beowulf cluster was purchased at transtec AG (Tübingen – D).

A drumscanner type `PRIMESCAN` from Heidelberger Druckmaschinen AG (Heidelberg – D) was used for digitalisation of the micrographs. The drumscanner was operated from a PowerMac 64 with a 400 MHz processor and 512 MB RAM.

2.1.7 Software

IMAGIC-5 The `IMAGIC-5` software package is a modular scientific image processing environment for the analysis of images in various subjects such as biology, medicine, and environmental sciences ([van Heel and Keegstra, 1981], [van Heel *et al.*, 1996]).

The various modules perform different procedures, for example Fourier transforms, alignments, multivariate statistical analysis, classification, angular reconstitution, and three-dimensional reconstruction from projections (tomography). The

modules will be explained in detail elsewhere. The IMAGIC-5 software was purchased at Image Science Software GmbH (Berlin – D).

EMAN Like IMAGIC-5, EMAN 1.7 is a modular software package specifically designed to make single particle reconstructions ([Ludtke *et al.*, 1999]). This software is a freeware and the version EMAN was acquired at the website http://ncmi.bcm.tmc.edu/ncmi/software/.

Linocolor The Linocolor 6.0.12 is a scanning software especially designed to operate the drumscanner and able to scan micrographs at very high resolution. The software was purchased along with the drumscanner.

CTFFIND3 CTFFIND3 is a program to determine the defocus and astigmatism from images of untilted specimens ([Mindell and Grigorieff, 2003]). The software is a freeware and was acquired at the website http://emlab.rose2.brandeis.edu/grigorieff/downloads.html.

ClustalX ClustalX 1.8 is a multiple sequence alignment program ([Higgins and Sharp, 1988], [Thompson *et al.*, 1997]). It provides an integrated environment for performing multiple sequence and profile alignments and analysing the results. ClustalX is a freeware and was obtained at the ftp-site ftp://ftp-igbmc.u-strasbg.fr/pub/clustalx.

MODELLER MODELLER 8v2 is a software for homology or comparative modelling of three-dimensional protein structures. With an alignment of a sequence to be modeled with known related structures, MODELLER automatically calculates a model containing all non-hydrogen atoms by satisfaction of spatial restraints ([Sali and Blundell, 1993]). MODELLER is a freeware and was acquired at the website http://salilab.org/modeller/download_installation.html.

PROCHECK The PROCHECK v.3.5.4 suite of programs provides a detailed check on the stereochemistry of a protein structure [Laskowski *et al.*, 1993]). PROCHECK is a freeware and was acquired at the website http://www.biochem.ucl.ac.uk/∼roman/procheck/procheck.html.

MOLREP MOLREP 9.3 is an automated program for molecular replacement and is part of the CCP4 programs suite ([Vagin and Teplyakov, 1997]). MOLREP is a freeware and was downloaded from the website http://www.ysbl.york.ac.uk/∼alexei/molrep.html.

Situs `Situs` is a package for docking crystal structures into low-resolution densitiy maps from EM ([Wriggers *et al.*, 1999]. The `Situs` package is freeware and was downloaded from the website `http://situs.biomachina.org/`.

Amira – ResolveRT `Amira - ResolveRT 4.0` is a professional general-purpose visualization and 3D-reconstruction software. 3D-volumes of different formats are displayed as polygonal surface models as well as tetrahedral grids. The software was downloaded from the site `http://www.tgs.com` and the license was purchased at TGS Europe (Düsseldorf – D).

Chimera `Chimera Version 1 Build 2199` is a highly extensible, interactive molecular graphics program ([Pettersen *et al.*, 2004]). `Chimera` is a freeware and was obtained at the website `http://www.cgl.ucsf.edu/chimera/download.html`.

Operating systems: SuSE LINUX – Debian LINUX The cluster used for the initial calculations was run on `SuSE LINUX` operation system versions 9.0 to 10.0. The 'Beowulf' cluster was run on the `Debian 3.1` free operating system.

Operating system: MacOS The operating system on the PowerMac 64 was OS D1-9.2.2 and was purchased together with the hardware.

2.2 Biochemical methods

2.2.1 Hemocyanin purification

The hemolymph was withdrawn from the dorsal sinus of a living *Limulus polyphemus* with a syringe and centrifugated for 30 min at 3,300 g in an `Eppendorf` centrifuge. The cell free supernatant was then centrifugated for two hours at 130,000 g in an air driven `Airfuge Ultracentrifuge` from Beckman (Munich – D). The hemocyanin pellets were resuspended in a low-salt buffer with 100 mM Tris–HCl (pH 7.8), 10 mM $CaCl_2$, 10 mM $MgCl_2$.

2.2.2 Hexamerin purification

Hemolymph extraction Hemolymph was extracted out of third-instar larvae. In this stadium larvae have a extraordinary high concentration of hexamerins in the hemolymph. The selected larvae were washed in cooled PBS buffer, narcotized with ice, and the skin torn open with precision tweezers from Roth (Karlsruhe – D) in a droplet of PBS-buffer. Five larvae were prepared as described in 10 μl droplets, which were then collected in caps for further processing and analysis.

PBS-buffer	137 mM NaCl 5 mM KCl 8 mM Na_2HPO_4 1.5 mM KH_2PO_4 pH 7.2

Native polyacrylamidgelelectrophoresis The hexamerins were purificated according to [Mousseron-Grall *et al.*, 1997]. In the native polyacrylamidgelelectrophoresis the hexamerins lost their quaternary structure maintaining the ternary structure of the subunits. The separation of the three subunit types (α, β, and γ) ocurred in an electric field according to their molecular mass, charge and ternary structure.

The extracted hemolymph was diluted with sample buffer in a ratio of 1:5 and applied to a 160×120×1 mm gel. The separation was performed with 15 mA, maximal voltage, and 4°C for approximately three hours. The left and right outer lanes were cut off the gel and stained with coomassie to determine the position of the hexamerins in the unstained gel. These positions were cut out of the unstained gel and transferred with 2 – 3 ml electroelution buffer into dialysis tubes from Roth (Karlsruhe – D).

Protogel	37.5:1 acrylamid solution (30% acrylamid, 0.8% bisacrylamid)
Polymerisation starter	10% APS TEMED
Elecrophoresis buffer	4% Natriumtetraborat ($Na_2B_4O_7$)
5.6% Gel	110 mM Tris/Citrat pH 9
Sample buffer	120 mM Tris/Citrat 8% Saccharose Bromphenolblue pH 9
Coomassie solution	0.2% Coomassie R 250 40% Methanol 7.5% Acetic acid
Unstaining solution	20% Isopropanol 7% Acetic acid

Electroelution To extract the separated native subunits out of the gel, the samples were eluted in an electric field. The dialysis tubes were brought into a horizontal gelelectrophoresis chamber filled with electroelution buffer and elution was performed at 20 V, maximal amperage, and 4°C for approximately twelve hours. Prior to the extraction of the samples, a high inverse voltage was applied to release subunits that were possibly attached to the membrane of the dialysis tube.

Electroelution buffer	50 mM Tris
	38 mM Glycin

Dialysis During the native gel electrophoresis, the hexamerins loose their quaternary structure and dissociate into their (native) subunits. To reasemble hexamers from the subunits, the dialysis tubes from the electroelution step were transferred into a reassociation buffer. The tubes were dialysed for approximately twelve hours at 4°C. After that, the samples were transfered from the dialysis tubes into caps.

Reassociation buffer	15 mM KH_2PO_4
	15 mM Na_2HPO_4
	10 mM $MgCl_2$
	140 mM NaCl
	pH 6.3

Protein enrichment Due to the fact that the samples gained through the purification had a rather low concentration of less then 0.1 mg/ml, a protein enrichment was necessary. A tenfold enrichment was achieved using YM-100 Microcon centrifugal filters from Millipore (Bedford) with membranes of a threshold of 100,000 Da.

Western-Blot Proteins separated in the electrophoresis were transfered onto a membrane of nitrocellulose to be detected by immunoblotting. The transfer occures with the application of an electric field to the gel so that the separated proteins transfer through the gel onto the membrane maintanining the same pattern as they had in the electrophoresis. Hydrophobic and charged interactions bind the proteins to the membrane.

The system used is called 'semi-dry'. Gel and membrane were placed between eighteen layers of Whatman filter papers soaked with transfer buffer in a blotting chamber SD1 from cti (Idstein – D). The transfer was performed at 0.8 mA/cm^2 for two hours. After the transfer, the membrane was reversibly stained with a diluted Ponceau S solution from Sigma (Deisenhofen – D) for approximately two minutes. The membrane was cut into five strips, two for each primary antibody in different concentrations and one control.

The strips were then washed briefly with H_2O_{dd} and incubated in blocking buffer for one hour. This step was taken to prevent non-specific interactions between the

antibodies and the membrane. After the blocking, four strips were incubated with the corresponding primary antibody overnight at 4°C under gentle agitation. Primary antibody solutions were diluted in proportions of 1:5,000 and 1:20,000 with blocking buffer. After the first incubation, the strips were rinsed three times with TBS-T to remove unbound primary antibody. The four strips incubated with the primary antibodies and the control strip were then incubated with the secondary antibody diluted in a proportion of 1:7,500 in TBS-T for two hours, at room temperature, and gentle agitation. At the end, the strips were rinsed again.

Detection was performed with the colorimetric detection method. Hereby a soluble dye is converted into an insoluble form of a different colour precipitating next to the enzyme staining the membrane. The strips were incubated with AP buffer for fifteen minutes in the dark. Incubation was stopped by rinsing the strips with H_2O_{dd} and drying them with Whatman filter paper.

Transfer buffer	39 mM Glycin 48 mM Tris 0.037% SDS 20% Methanol pH 8.3
10x TBS	100 mM Tris 1.5 M NaCl pH 7.5
TBS-T	0.3% Tween 20 in 1x TBS
Blocking buffer	5% (w/v) non-fat dry milk in TBS-T
10x Developing buffer	100 mM Tris 1 M NaCl 50 mM $MgCl_2$ pH 9.3
AP buffer	10 ml 1x developing buffer 33 μl BCIP (50 mg/ml in 100% Dimethylformamid) 66 μl NBT (50 mg/ml in 70% Dimethylformamid)

2.3 Electron microscopy

2.3.1 Negative staining

The negative staining technique was used to analyse the reasembling process of the hexamerin ([Harris, 1997], [Harris, 1999]). A 5 μl droplet of a 0.1 mg/ml sample was spread onto a holey carbon coated grid that had previously been glow-discharged. After approximately 10 seconds the exceeding volume was blotted with filter paper and the grid washed consecutively in four 10 μl droplets of distilled water. The grid was then covered with a 10 μl droplet of stain (2% uranyl acetate with 0.1 mM octyl-β-D-glycopyranosid) and after approximately 5 seconds blotted dry with filter paper.

2.3.2 Cryo-plunging

Cryo-plunging was performed according to [Adrian *et al.*, 1984]: a droplet of the purified protein (hemocyanin or hexamerin) with concentrations between 0.1 and 0.3 mg/ml was applied to a glow-discharged holey carbon support film. After removing exceeding fluid by blotting, the grid was rapidly plunged frozen in liquid ethane.

The plunging for the hemocyanin under controlled atmosphere was performed in a 'plunging tube' (Figure 2.2). A constant stream of gas inside the tube provided stable and controlled conditions during plunging. For the oxygenated state of the hemocyanin, a gas stream with a mixture of 25% O_2 and 75% N_2 was used.

2.3.3 Cryo-electron microscopy

The hemocyanin cryo-transfer was performed using a 626 single tilt cryotransfer system from Gatan (Gatan GmbH, Munich – D) to a Tecnai F30 equipped with a field emission gun (FEG) operated at an accelerating voltage of 200kV. Electron micrographs were recorded under low-dose conditions at 59,000x instrumental magnification. The underfocus was set between 1.0 μm and 5.0 μm. The Agfa Scienta negatives were developed for twelve minutes in full-strength Kodak D19 developer.

The LSP-2 protein cryo-transfer was performed using a Gatan model 626 DH cryo-holder to a Philips CM200 equipped with a field emission gun (FEG) operated at an accelerating voltage of 200kV. Electron micrographs were recorded under low-dose conditions at 50,000x instrumental magnification. The underfocus was set between 0.77 μm and 1.5 μm. The Agfa Scienta negatives were developed for twelve minutes in full-strength Kodak D19 developer.

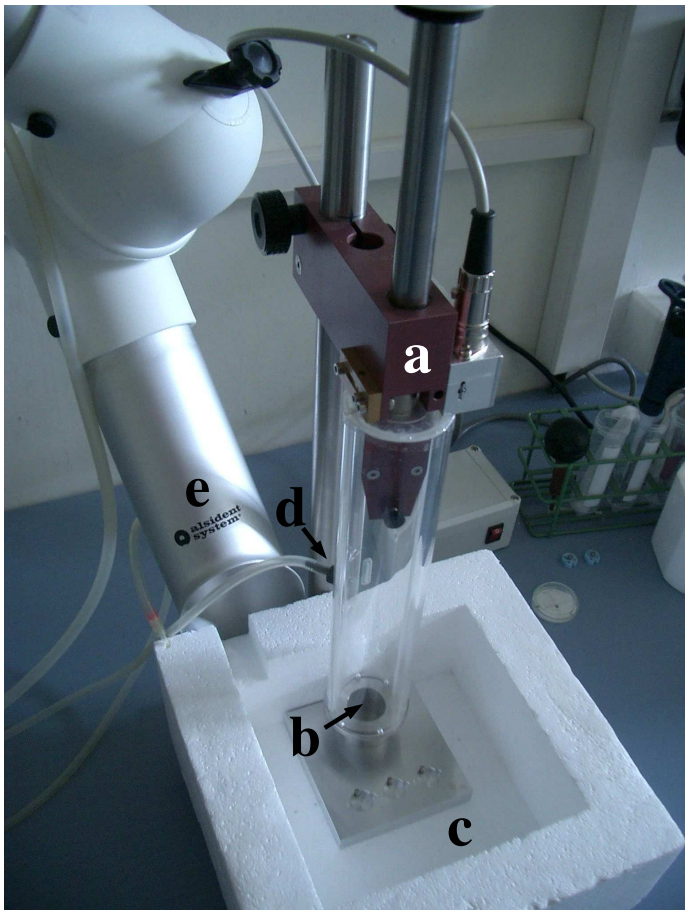


Figure 2.2: Plunging tube. **a** – guillotine, **b** – reservoir for liquid ethane, **c** – reservoir for liquid nitrogen, **d** – gas inlet, **e** – exhaustor.

2.4 Bioinformatical methods

2.4.1 Digitisation

Drift- and astigmatism-free electron micrographs were scanned and digitized using a `PRIMESCAN` drumscanner from Heidelberger Druckmaschinen AG (Heidelberg – D). The hemocyanin micrographs were scanned at 8 bit and with 10.6 and 5.9 μm step size with a pixel corresponding to 1.8 and 1.0 Å, respectively, on the micrograph. LSP-2 micrographs were scanned with 10.6 μm step size with a pixel corresponding to 2.12 Å on the micrograph. The scanned files were saved in the `tiff`-format

2.4.2 Particle selection

Images of single particles were selected semi-automatically with the module `boxer` from the `EMAN 1.7` software package. Focal pair micrographs processing was performed with this module as well.

2.4.3 Preprocessing

CTF correction The image produced by an electron microscope is the result of the interaction between the electron beam and the analysed sample. The image is formed by a combination of elastic and inelastic electron scattering induced by the sample. The importance of the inelastic scattering can be neglected in the context of the single particle analysis, once it produces an almost featureless background at low resolution and falls of towards higher resolution ([Zhu *et al.*, 1997]).

Most of the structural information of the sample is thus contained in the elastic scattered electrons, that produce the phase contrast of unstained specimens. The phase contrast is modulated by the contrast transfer function (CTF) of the microscope: it is a function of defocus, astigmatism, spherical aberration, and electron wave length.

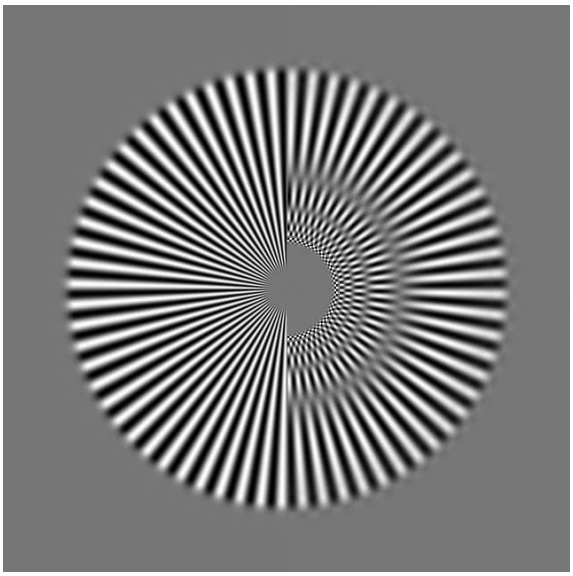


Figure 2.3: Siemens star with imposed CTF. The effect of the phase contrast on an image is clearly visible in the skipping of the black an white colours of the spikes on the right half of the Figure.

The CTF correction is of crucial importance to achieve highly resolved cryo-EM structures. The CTF of the microscope causes the so called 'Thon rings' ([Thon, 1966]) that can be observed in a power spectrum and are amplitude modulations and phase reversals of an image. During correction, the phases of the negative lobes of the CTF are 'flipped' and therefore, the precise determination of defocus and astigmatism are of great importance. The effect of the phase contrast on an image can be seen in Figure 2.3, a siemens star. The right half of the figure has an imposed CTF: the colour of the spikes skips from black to white (and vice-versa) when passing through the grey concentric rings. Latter are the regions with zero information.

The estimation of the defocus and astigmatism parameters were performed with the program `CTFFIND3` and the CTF correction was realized with the module `transfer` from the `IMAGIC-5` software package.

Band-pass filtering The CTF-corrected images contain information (spatial frequencies) that do not represent the analysed macromolecules. Images are thus filtered to restrict the information to the relevant spatial frequencies.

The disturbing low spatial frequencies, which are associated with density ramps or other large-scale effects, can seriously interfere with alignment procedures that use cross correlation functions. On the other hand, high spatial frequencies beyond the expected maximum resolution mainly represent noise and hence can be suppresed.

The applied band-pass filter is the product of a wide Gaussian low-pass filter and a narrow Gaussian high-pass filter. The low-pass filter cuts away the high spatial frequencies whereas the high-pass filter cuts away the low spatial frequencies. Band-pass filtering was executed with the module `incore-prepare-filtered` from the `IMAGIC-5` software package. The cutoffs where generally determined as follows:

High-pass cutoff	$2 \times$ pixel size (Å) $\div$ particle size (Å)
Low-pass cutoff	$2 \times$ pixel size (Å) $\div$ maximal expected resolution (Å)

In later stages of the reconstruction, the cutoffs were gradually shifted to the high spatial frequencies.

After filtering, a circular mask was applied to all images to remove unwanted background. The images were normalised by floating the average density of the greyvalues to zero and by setting an arbitrary variance value of three within the mask.

2.4.4 Reference-free alignment

To approach the structure of the studied molecules without biasing the data set right from the beginning (thus avoiding 'self-fulfilling prophecy processing', [van Heel *et al.*, 2000]), a 'reference free' alignment procedure was applied. Images of a data set were translationally centred according to their rotationally averaged total sum, maintaining the original rotational orientation of every single image. The following multivariate statistical classifcation and image averaging (see below) identified similar images with comparable orientations and produced improved views. These improved views were then used to start an unbiased multiple reference alignment ([van Heel and Stöffler-Meilicke, 1985], [Dube *et al.*, 1993], [van Heel *et al.*, 1996]).

The reference free alignment was performed with the module `align-mass` from the `IMAGIC-5` software package.

2.4.5 MRA – multiple reference alignment

The aim of the mutliple-reference alignment is to align a mixed population of noisy molecular projection images with respect to a large number of references. A precise alignment is crucial for averaging the projections (see below) and thus for improving of the signal-to-noise ratio (SNR).

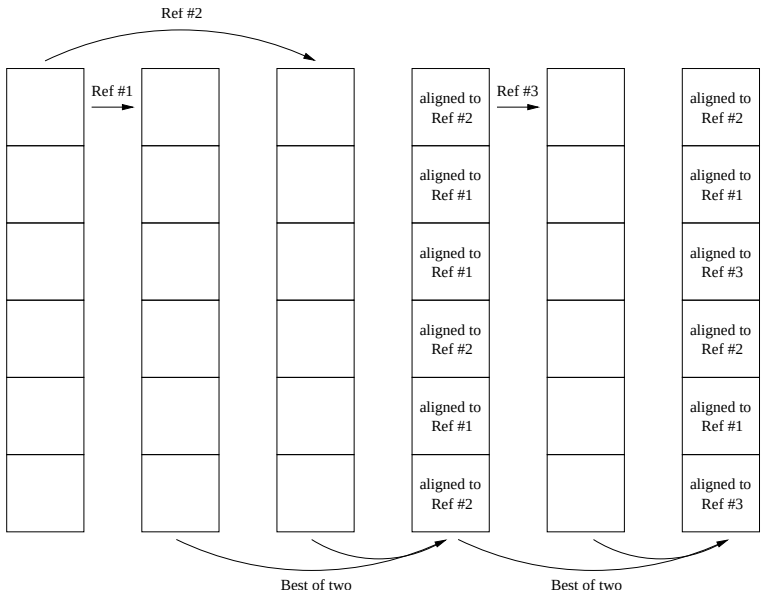


Figure 2.4: Procedure of the MRA - multiple reference alignment according to [van Heel *et al.*, 2000] (modified). Images are aligned in respect to the reference with the best resulting correlation coefficient.

In the first step of an MRA, each image of the projection images file stack is aligned to reference no. 1 and a correlation coefficient is associated. In the subsequential step, images are aligned to reference no. 2 and again a correlation coefficient is associated. The criteria to decide which alignment is kept (with respect to reference no. 1 or no. 2) is the correlation coefficient. The result of these first two steps is a mixed data set of differently aligned images. In the next step, the mixed data set is aligned to reference no. 3 and again alignments with the the higher correlation coefficients are kept. The procedure is repeated for all references. MRA procedures need most of the computing resources during a single-particle analysis.

During an alignment a projection image can be rotated or translated in respect to its reference. Rotation and translation are *per se* independent from each other, but most commonly both are used iteratively to optimize an alignment.

Translational alignment In a translational alignment, the projection image and the reference are initially 2D-fourier transformed. The fourier transforms are then multiplicated and the result, a cross-correlation function between both images, is 2D back-transformed from fourier space into real space. The resulting image in real

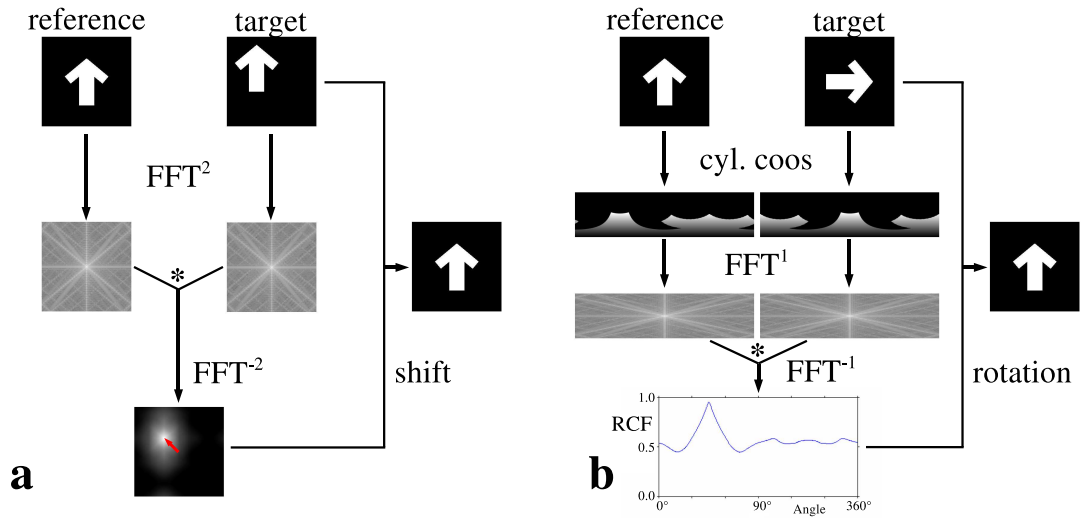


Figure 2.5: Scheme of the alignment of a target against its reference according to [Schatz, 1992] (modified). **a** – Translational alignment. Reference and target are 2D Fourier-transformed (FFT^2), the results multiplicated ($*$), and transformed back into real-space (FFT^{-2}). The resulting image has a strong peak that is shiftet by a vector $\vec{r}$ (red arrow) from the centre of the image. The target is finally shifted according to this vector. **b** – Rotational alignment. Reference and target are represented in a cylindrical coordinate system (cyl. coos), 1D Fourier-transformed (FFT^1), the results multiplicated ($*$), and transformed back into real-space (FFT^{-1}). The resulting RCF has a maxima at the angle by which the target is rotated in relation to the reference. The target is finally rotated according to this angle.

space has a strong and easily detectable peak that is shifted in relation to the image centre. The shift of the peak corresponds to the shift of the projection image in relation to its reference and can be described as vector $\vec{r}$. The projection image is then shifted according to this vector (Figure 2.5 a).

Rotational alignment For a translational alignment, projection image and reference are initially represented in cylindrical coordinates to then be 1D-fourier transformed. The fourier transforms are multiplicated and the result, this time a rotational correlation function (RCF), is 1D back-transformed. The result in real space is the RCF plotted against the rotational angle. The highest peak of the RCF correlates with the angle which the projection image is rotated in relation to its reference. The projection image is then rotated according to this angle (Figure 2.5 b).

The multiple reference alignment (translational and rotational) was accomplished with the module `multi-reference-alignment` from the `IMAGIC-5` software package.

2.4.6 MSA – multivariate statistical analysis

After a reference-free alignment or once images were aligned to references, a pattern-recognition procedure was used to identify the different views of the analysed macromolecules. The technique applied was the MSA in the form of the *eigenvector-eigenvalue* analysis ([van Heel and Frank, 1981], [Lebart *et al.*, 1984]).

In a first step, aligned images with $n \times n = m$ pixels were represented as single points in a m-dimesional 'hyper'-space. Every pixel represents a vector in this hyperspace and the position of a single point is determined by the linear combination of the pixel densities. Thus, a single point in the MSA-hyperspace is entirely equivalent to its 2D molecular image. A stack of images will result in a cloud of points in the hyperspace, in which similar images will end in close positions whereas different images will be far apart.

Then, the original coordinate system in the hyperspace is rotated in such a way that the new coordinate system is optimally adapted to the shape of the cloud produced by the stack of images. With the modulation distance metrics used here ([Borland and van Heel, 1990]), the first axis of the new coordinate system corresponds to the centre of the cloud and describes the total average of the aligned images. The second axis, orthogonal to the first, describes the most important direction of the differences within the stack of images. The third axis, orthogonal to the first two, then describes the next most important direction of the remaining differences, etc. The significance of the inter-image variance (*eigenvalues*) described by every new axis (*eigenvectors*) falls and the higher axes represent noise rather than structure information. Therefore, the comparison between images can be reduced to the first (max. 69) *eigenvectors*. This reduction facilitates the understanding of the data set and minimizes computational requirements.

To restrict the MSA to the region with the relevant macromolecule information, a narrow circular mask was applied to the images to reduce the influence of the background.

The MSA procedure and the creation of the MSA-mask were performed with the modules `msa-run` and `test-image`, respectively, from the `IMAGIC-5` software package.

2.4.7 MSA classification

The MSA processed data sets were then classified with an automated hierarchical ascendant classification (HAC) combined with a moving-elements-refinement ([van Heel, 1984a], [Borland and van Heel, 1990]). The aim of this classification was to minimise the intra-class variance, maximising at the same time the inter-class variance between the centres of mass of the class averages.

The HAC begins considering every image of the stack as a single class, which are then succesively merged, two at a time, following the criterium of the lowest possible increase of the intra-class variance (Ward criterium, [Ward, 1982]) at the

corresponding level of the procedure. Classes are merged until the whole data set is grouped into one single class. The grouping process can be tracked in a 'classification tree' (Figure 2.6). The number of classes is chosen by the user and the classification tree is cut at the level that corresponds the requested number.

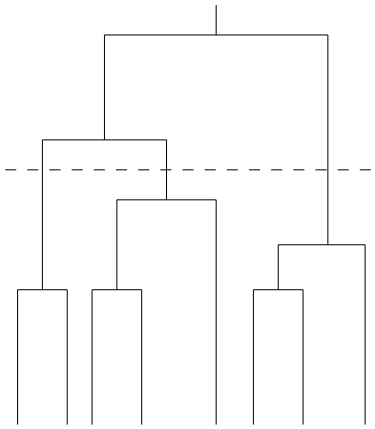


Figure 2.6: Scheme of the hierarchical ascendent classification according to [Schatz, 1992] (modified). At the beginning (bottom), every image of the stack is a single class. The classes are gradually merged, two at a time, following the Ward criterium. The dotted line represents the level of the requested number of classes.

The primary clustering obtained with the HAC is then refined by allowing all members of the classes to move to other classes (moving-elements-refinement) if this transfer leads to a reduction of the variance ([van Heel, 1984a], [van Heel, 1989]). The refinement process is performed iteratively until the total grouping stabilizes. Finally, all images belonging to a class are averaged, improving the SNR. The SNR is proportional to the number of averaged images. Class averages with their improved SNR may be used as references for new MRA/MSA/HAC iteration rounds (see Subsection 2.4.10 below) and are of great importance for a precise determination of the Euler angles.

The MSA classification was realised with the module `msa-classify`, class averging with the module `msa-sum`, both from the `IMAGIC-5` software package.

2.4.8 Angular reconstitution

Two 2D projections of the same 3D object always have a 1D line projection in common (common line projection theorem, [van Heel, 1987]). The common line projection theorem is the real-space equivalent of the Fourier-space common lines theorem ([DeRosier and Klug, 1968]). The relative Euler angles orientations of the 2D projections to one another can be determined through the angles between such

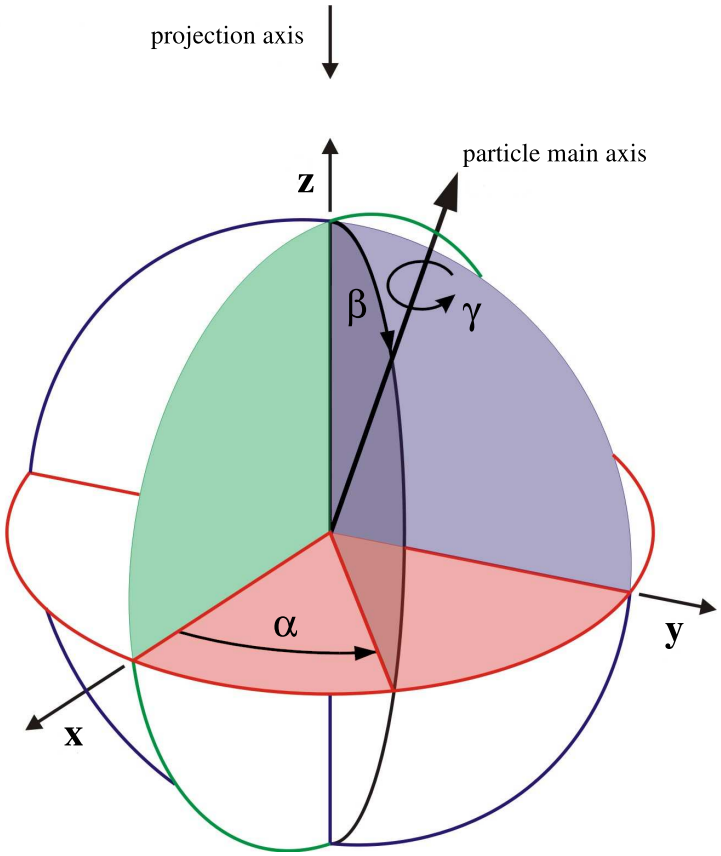


Figure 2.7: Definition of the Euler angles in a common coordinate system according to [Stohr, 2001] (modified). The α-angle is defined as the in-plane rotation perpendicular to the projection axis Z; the β-angle is the angle of the tilting of the particle main axis; the γ-angle represents a rotation around the particle main axis. When the particle main axis coincides with the projection axis Z, then the α-angle and the γ-angle rotate in the same plane.

common line projections. The definition of the Euler angles is shown in Figure 2.7. For macromolecules with no structural symmetry (i.e., a ribosome), at least three different projections are required to solve the orientation problem. The first two 2D projections can freely rotate around their common line (Figure 2.8). The relative orientation of these first two projections is finally determined through the third projection that has different common lines with each of the initial projections.

Finding the orientational relationships between the 2D class averages produced by the MSA/HAC produres, allows the reconstruction of the analysed 3D-structure. The angular reconstitution technique is thus *a posteriori* method to find relative orientations of 2D projections.

To find the relative orientations, first 'sinograms' of the class averages were created. A sinogram is a stack of the 1D line projections of a 2D projection. Then, the sinograms of two 2D projections were compared line-by-line in a sinogram correlation function (Figure 2.9). The sinogram correlation function will have a maximum

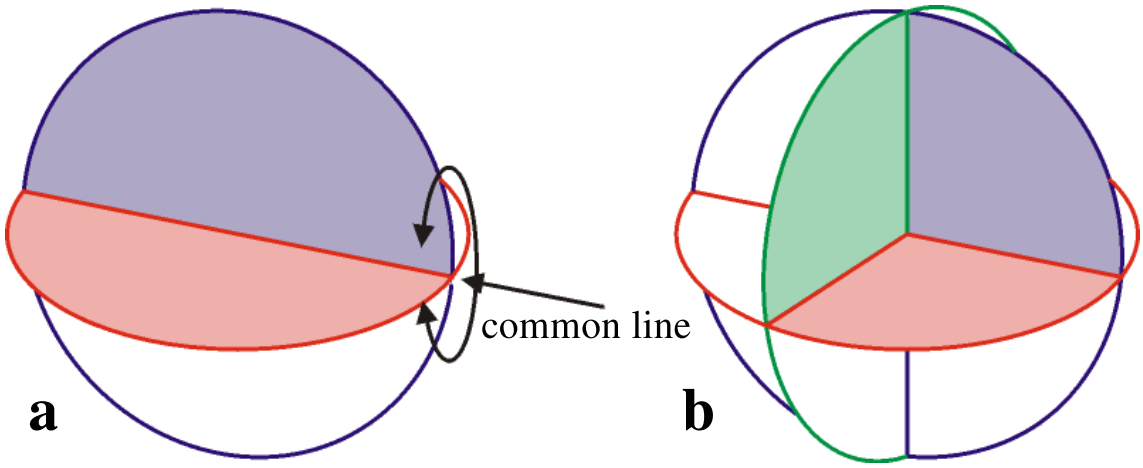


Figure 2.8: 2D projections from the same 3D object represented in the Fourier space according to [Stohr, 2001] (modified). **a** – Two 2D projections of the same 3D object share one line in common, around which they can freely rotate. **b** – With a third 2D projection, the rotation around the first common line is fixed and the relative orientations of all 2D projections defined.

at the position corresponding to a pair of shared line projections. The higher the symmetry of a particle, the more peaks the sinogram correlation function will have.

Due to the fact that correlation functions and correlation coefficients are 'squared' similarity measures, low frequency components of the data may be overweighted during the search of the relative orientations ([van Heel *et al.*, 1992a]). To prevent this, the sinogram lines are 'amplitude square root filtered' ([van Heel *et al.*, 1992b]) and are normalised prior the sinogram correlation.

The introduction of a new class average to a set of projections with already determined relative orientations is performed over all Euler-angles orientations corresponding to the full 'asymmetric triangle' (see below) for the given point-group symmetry ([Schatz *et al.*, 1995]). The normalised standard deviation of the peak heights among all corresponding peaks serves as an internal consistency check.

The asymmetric triangle of an asymmetric particle like the ribosome covers the entire unit cell. For a particle with a D3 symmetry, i.e., a hexamerin, the asymmetric triangle reaches from the 'north pole' of the unit sphere to the equator and there it spans $360°/3 = 120°$. The hemocyanin from *Limulus polyphemus* with a D2 symmetry spans $360°/2 = 180°$ along the equator.

The first cryo-EM structure can be used to create a set of images by reprojection within the asymmetric triangle ([Orlova and van Heel, 1994], [Schatz *et al.*, 1995], [Serysheva *et al.*, 1995]). This set, also called 'anchor set', is consistent in terms of relative orientations and will contain less noise than the original class averages from which the cryo-EM structure was generated. The search of Euler-angles in later phases of the analysis is performed by comparing each of the class averages with the anchor set.

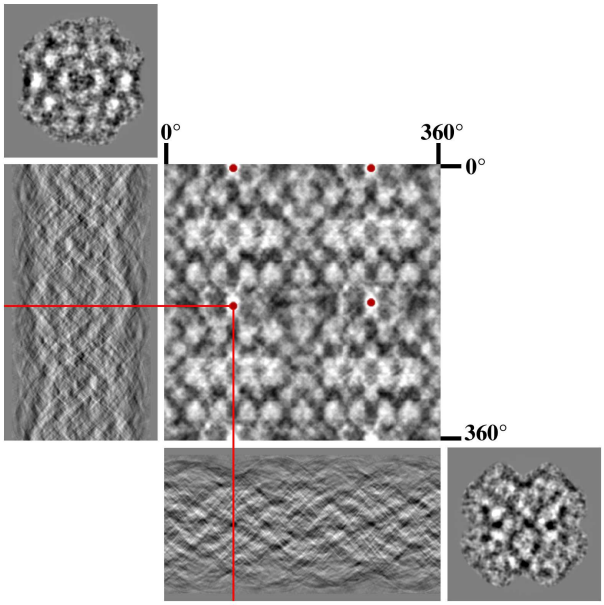


Figure 2.9: Sinogram correlation function according to [van Heel *et al.*, 2000] (modified). 2D class averages (small squares) are 1D projected leading to sinograms (rectangles). Sinograms are compared line-by-line in the sinogram correlation function (big square), in which every point is the correlation coefficient of two lines of the sinograms. The D2 symmetry of the hemocyanin from *Limulus polyphemus* can be identified in the four correlation peaks (red dots) of the sinogram correlation function.

The search of the Euler-angles (relative orientations) of class averages was performed with the module `angular-reconstitution` from the `IMAGIC-5` software package.

2.4.9 Three-dimensional reconstruction

The mathematical principle of reconstructing a 3D object from a set of 2D projections was analytically solved by Radon ([Radon, 1917]). He proved that a n-dimensional function can be described by its *(n-1)*-dimensional projections. [DeRosier and Klug, 1968] presented the 'projection theorem' and introduced a technique to extract the 3D information from electron micrographs: a 2D projection of a 3D object corresponds to a central section of the 3D Fourier transform of the object. The 3D-reconstruction is performed by filling the Fourier space with central sections derived from the 2D projections.

The central section of a 3D object in Fourier space is not infinitely thin as the mathematical plane in the analytical concept of Radon, but a 'slab' (Figure 2.10) with the thickness of $\sim 1/D$ ([Harauz and van Heel, 1986a], [Harauz and van Heel, 1986b]). D is the the largest linear dimension of the 3D

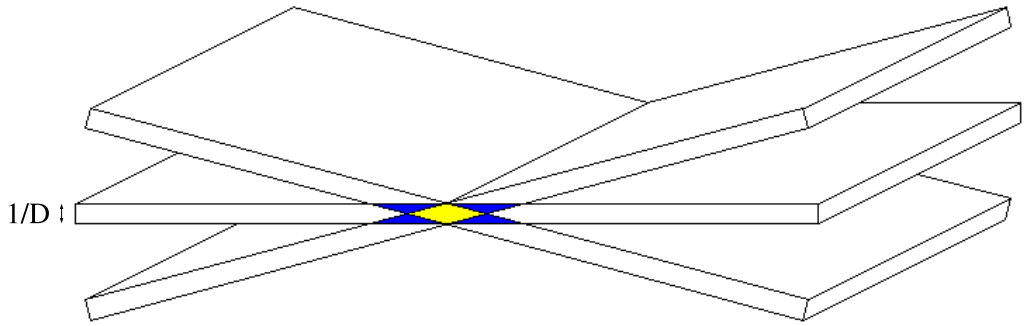


Figure 2.10: Overlapping central sections in Fourier space according to [van Heel *et al.*, 2000] (modified). The slabs have a thickness of 1/D, where D is the largest linear dimension of the 3D object. Overlapping areas close to the origin (that correspond to low spatial frequencies) are coloured in blue and yellow.

object. Because the slabs have a finite thickness, they overlap over more then one single line in Fourier space (blue and yellow areas in Figure 2.10). The closer to the origin of Fourier space, the more the slabs overlap (yellow area). Therefore, simply adding the data in Fourier space to perform a 3D-reconstruction would end in overweighting the low spatial frequencies that are close to the origin. To avoid this overweighting, exact-filter algorithms were introduced ([Harauz and van Heel, 1986b], [Radermacher, 1988]) in which a specific filter is computed for each projection in the reconstruction.

The module `true-threed-reconstruction` from the `IMAGIC-5` software package was used to perform the 3D-reconstruction.

2.4.10 Iterative refinements

The whole process of the single particle analysis is a constant refinement procedure at different levels. Beginning with the sample preparation (Figure 2.11 top left), the cryoelectron microscopy (see Subsection 2.3.2) can serve as a purity check for the sample. Once the sample preparation is optimised, a large number of micrographs is generated, subsequently digitised at high resolution (see Subsection 2.4.1), and particles selected (see Subsection 2.4.2). In the first step of the reconstruction, class averages obtained through the reference-free alignment (see Subsection 2.4.4) are used as references for the first MRA (see Subsection 2.4.5). After stabilisation of the class averages, the Euler angles are assigned (see Subsection 2.4.8) and a first cryo-EM structure is calculated (see Subsection 2.4.9). Few reprojections of the first 3D recontruction are used for the refinement of the Euler angles of the class averages. When the cryo-EM structure gets stable, many reprojections are generated for the next round that starts with a new MRA. After some refinement rounds and the final stabilisation of the cryo-EM structure, the structure interpretation process starts.

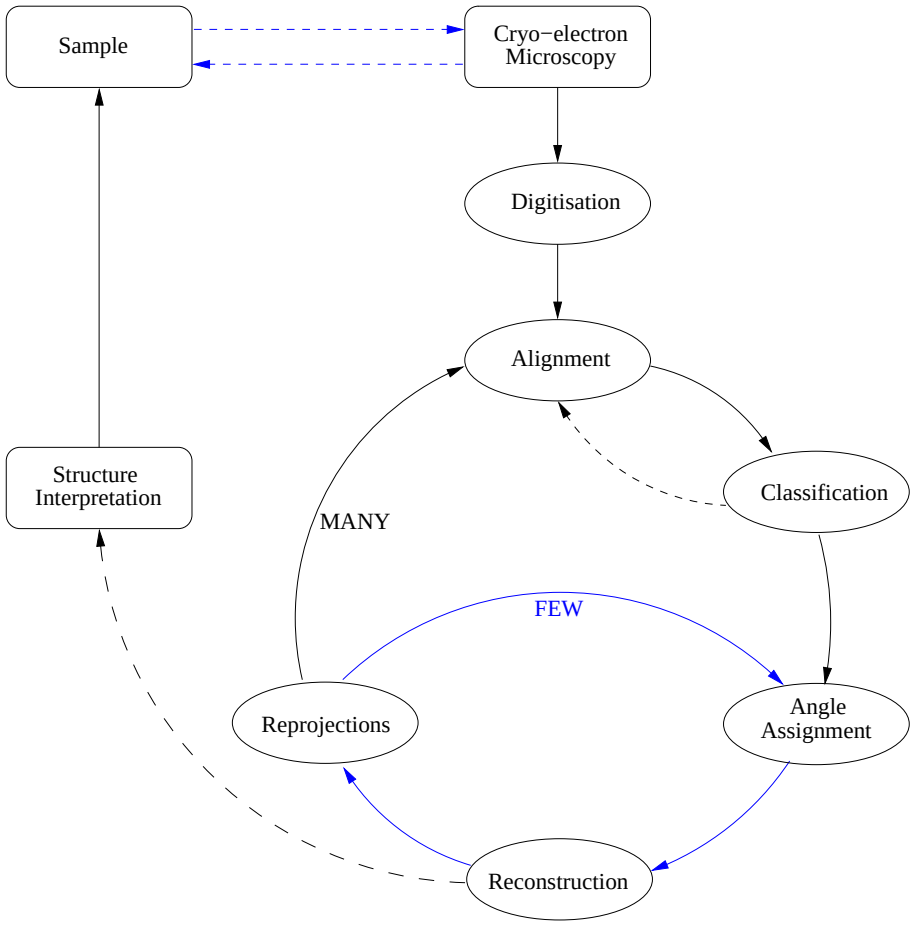


Figure 2.11: Single particle analysis according to [van Heel *et al.*, 2000] (modified). The quality of the purified sample is first assesed with the cryo-electron microscopy (top left – the explanation proceeds clockwise) and if necessary optimised. A large number of micrographs is generated and subsequently digitised at high resolution. Single particles are extracted from the micrographs, CTF corrected, and pretreated to then be aligned (MRA). Class averages generated by MSA classification can serve as references for new rounds of MRA or Euler angles are assigned to former. Class averages with assigned Euler angles are used to calculate a cryo-EM structure. Once a cryo-EM structure is calculated, reprojections are generated and used as references for the Euler angles assignment or for new rounds of MRA. The process is iterated until the results are stable. A stable cryo-EM structure allows structure interpretations.

2.4.11 Projection matching

After achieving a median resolution (∼15 – 10 Å), the technique of projection matching was used to improve the cryo-EM structure. To do so, I first adapted the routine of 3D-reconstruction from the `EMAN` software package to our `IMAGIC-5` environment. Michael Stohr (Institute of Zoology – University of Mainz) developed a C-programme called `create-mra-classes` upon this procedure. The programme combines modules from the `IMAGIC5` software package and shell scripts to generate class averages based on Fourier correlation values.

A data set is aligned against a multiple reference file with a large number of images (> 1,000 references) that preferably covers the asymmetric triangle uniformly. Each image of the data set is aligned to the refence with the highest resulting cross-correlation coefficient (see Subsection 2.4.5). Images of the data set, which were aligned against the same reference, are then averaged and the Euler angles of latter are assigned. Thus, the creation of the class averages is not based on MSA and HAC, and Euler angles assignment is not based on the common line theorem, but on the Fourier correlation coefficient.

2.4.12 Molecular modelling

Molecular modelling was performed with the `MODELLER` software. The alignments of the sequences for the modelling were calculated with the `ClustalX` software.

The models of the subunit types II, IIIA, IV, and VI from *Limulus polyphemus* and the subunit types e, g, and f from *Eurypelma californicum* were built using the 3D-structure of the hemocyanin from *Limulus polyphemus* in the oxygenated state (RSCB PDB-IDCODE: 1NOL, [Magnus *et al.*, 1994]) as template for the homology structure modelling. For each sequence, 20 different models were genereated.

The model of the LSP-2 hexamer was built in coorperation with the Department of Molecular Biophysics (University of Mainz) using the three-dimensional structure of the hexameric hemocyanin from *Panulirus interruptus* (RSCB PDB-IDCODE: 1HCY, [Volbeda and Hol, 1989a]) as template for the homology structure modelling. For this protein, 31 hexameric models were generated.

2.4.13 Molecular fitting

In a first step, densities of the subunits from the hemocyanin were isolated from the whole EM density map with the `Resolve RT` software. To choose the best model of the 20 generated for each subunit type (from *Limulus* and *Eurypelma*), each model was fitted into its respective isolated subunit density with the `MOLREP` software. The model with the highest correlation coefficient of the 20 generated was then chosen.

The second step was to fit the chosen models of all subunit types into the whole EM density map. This work was done in cooperation with the Health Science Center – Houston – University of Texas with a reprogrammed version of the `Situs` software. Homo-48mers of each subunit type were generated and the final result was assembled by selecting the subunits in their respective positions within the 8×6mer with `Chimera`.

For the hexamerin, the hexameric models were fitted into the density map with the `Chimera` software. The model with the highest correlation coefficient was chosen.

2.4.14 Stereochemical quality of the models

The stereochemical quality of the generated models was checked with the `PROCHECK` software. Its outputs give an assessment of the overall quality of the structure as compared with well refined structures of the same resolution and also highlight regions that may need further investigation.

2.4.15 Structure predictions

For the hexamerin, two different structure predictions were made with servers. The prediction of the secondary structure of the loop with the insertion of ∼20 amino acids between β-sheets 3B and 3C (convention of [Gaykema *et al.*, 1986]) was calculated with the `PSIPRED` Protein Structure Prediction Server (`http://bioinf.cs.ucl.ac.uk/psipred/`) and the `SCRATCH` Servers (`http://www.igb.uci.edu/tools/scratch/`). The prediction of the *N*-glycosylation site was calculated with the `NetNGlyc 1.0` Server.

Chapter 3

Results and Discussion of the hexamerin LSP-2 from *Drosophila melanogaster*

3.1 Introductory remarks

For a better clarity, results and discussion of each protein are treated as a unit and not in separate chapters.

As shown in Figure 2.7 of the Subsection 2.4.8, the Z axis is defined as the main symmetry axis of the molecule and this is the axis along which, by the `IMAGIC-5` software package, the tomographic layers of the cryo-EM structure are stacked. In the following two chapters, all figures showing views along the main symmetry axes of the molecules (hemocyanin and hexamerin) are arranged in the order Z, X, Y.

3.2 Cryo-EM

A total number of 21 micrographs were recorded under low-dose conditions (Figure 3.1). Nominal defocus was set between 1.26 and 2.5 μm. With the scanning step size corresponding to 10.6 μm, the digitalised micrographs ended up with a size of 65 MB.

In the study of [Volbeda and Hol, 1989a], the terms 'trimer', 'tight dimer', and 'loose dimer' were introduced. [Hazes *et al.*, 1993] showed Figures in which the *trimer* was seen along the three-fold axis, the *tight dimer* was seen along the first two-fold axis, and the *loose dimer* was seen along the second two-fold axis. This convention is used throughout this chapter.

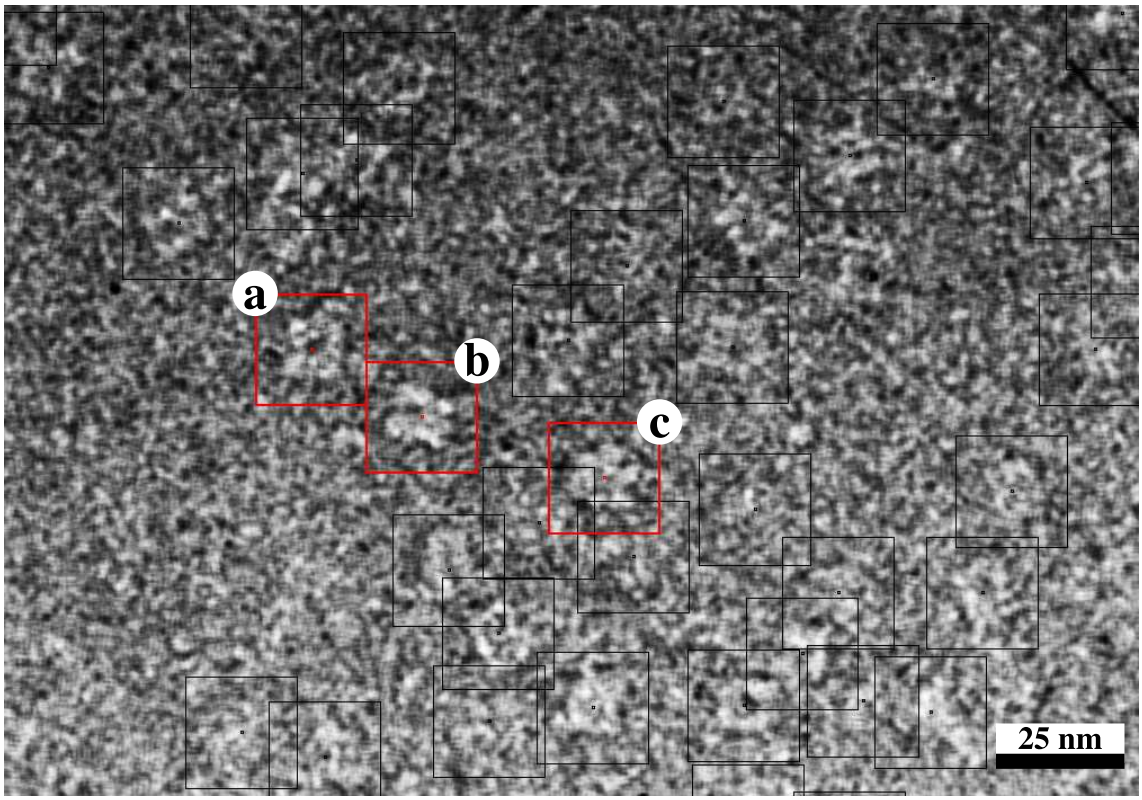


Figure 3.1: Cryo-EM negative of the *Drosophila melanogaster* hexamerin LSP-2. The Figure of the cryo-EM negative is a screenshot of the particle selection procedure with the programme `boxer`. The highlighted particles represent the views along the main symmetry axes. **a** – *tight dimer* view (almost a square, view along the first two-fold axis), **b** – *loose dimer* view (rectangular, view along second two-fold axis), **c** – *trimer* view (hexamer, view along the three-fold axis).

3.3 Single particle analysis

Single particles were semi-automatically extracted from the micrographs. The data set comprised 41,555 single particles. Defoci were estimated between 1.21 and 2.9 μm.

3.3.1 Angular reconstitution

The analysis started with a reference-free alignment of the 41,555 single particles. The resulting class averages (projections) from multivariate-statistical analysis (MSA) and hierarchical ascendent classification (HAC) were used for a first multi-reference alignment (MRA). After three alignment refinements, one top, two tilted, and three side view projections were selected. The criteria for the selection was the user based optical comparison with reprojections of the cryo-EM structure from *Palinurus elephas* ([Meissner *et al.*, 2003]). In a first step, the Euler angles of similar reprojections from *Palinurus* were assigned to the six selected projections

from *Drosophila*. Then, the angles of the six projections were manually varied until no further improvement could be achieved. These six projections were used as an 'anchor-set' for the rest of the data set. After angular reconstitution, a preliminary reconstruction was calculated. Re-projections of the first reconstruction were used for the next rounds of MRA and MSA.

After several refinement rounds, the data set was split according to classification and assigned Euler angles. The first subset of single particles comprised projections with a β-angle smaller than or equal to 75°, the second subset comprised single particles with a β-angle larger than 75°. The reason for the splitting was to improve the results of the MSA and HAC, which led to more detailed class averages. During refinement iterations, the data set was subsequently reduced by eliminating single particles of class averages with high cross correlation errors.

In the final refinement round of the angular reconstitution technique, the data set comprised 24,163 single particles divided in a subset with 13,309 particles with a β-angle smaller than or equal to 75° (top and tilted view projections), and a subset with 10,854 particles with a β-angle larger than 75° (side view projections). The top an tilted view projections were classified in 1,664 classes (∼8 particles per class), and the side view projections in 1,357 classes (∼8 particles per class). Of the 3,021 resulting class averages, 1,500 were used for the final cryo-EM structure with this technique. The final class averages were sorted by ascendant cross correlation error.

The resulting reconstruction of the angular reconstitution technique (Figure 3.2) achieved a resolution of 13.8 $Å_{1/2-bit}$ and 12.6 $Å_{3\sigma}$ as shown in Figure 3.3 a. The threshold was set to correspond to a molecular weight of 488 kDa. The distribution of the Euler angles is shown in Figure 3.3 b. The distribution of the projections in the asymmetric triangle is homogeneous with a light predomination of top and close-top-top/tilt images (β-angles smaller than 45°).

3.3.2 Projection matching

After the stabilisation of the cryo-EM structure by the technique of angular reconstitution, four rounds of projection matching were applied. The final cryo-EM structure was calculated with a selected data set of 11,292 single particles distributed between 3,880 class averages (average of ~ 3 particles per class). The first 88 class averages of the final cryo-EM structure (sorted by ascending cross correlation error) are shown in Figure 3.4. There is a clear predomination of side view projections (β-angle close to 90°). The first close to top view (projection along the Z axis) can be seen in position 34, and the first clearly tilted view (β-angle 18°) is found in position 67. Again, the threshold was set to correspond to a molecular weight of 488 kDa. The tomographic layers of the final cryo-EM structure are presented in Figure 3.5 and are visualised as stereo surface views in Figure 3.6.

The distribution of the Euler angles is shown in Figure 3.7 a. The homogeneous distribution of the angles is due to the technique of projection matching. As shown in Figure 3.7 b, the resolution reached 9.3 $Å_{1/2-bit}$ and 8.1 $Å_{3\sigma}$.

3.3.3 The adapted technique of projection matching

The projection matching routine I adapted from the `EMAN` software package to our `IMAGIC-5` environment (see Subsection 2.4.11 on page 29) yielded large improvements in the quality of the reconstruction after only a few steps. The LSP-2 project was the first in our group to profit from this new routine. This new method is discussed along with the hexamerin LSP-2 because the improvements can be elucidated more clearly on the level of the hexamer than on the complex level of the 8×6mer of the *Limulus* hemocyanin.

Figure 3.8 shows a comparison of the best result obtained by the technique of angular reconstitution after several iterative steps (left column) with the final result of projection matching after four rounds (right column). Hence, the improvement is obvious. In the left column, the cryo-EM structure already shows the structural features that will be discussed in detail in the following section, but the surface is smooth without namable details. The final cryo-EM structure in the right column resembles the overall shape of the left side, but reveals a significantly more detailed structure. The improvement of the quality assesed by the FSC confirms the optical impression: the resolution dropped from 13.8 Å to 9.3 Å (1/2-bit criteria) and 12.6 Å to 8.1 Å (3σ criteria), in both cases more than 4 Å after a few steps.

The new routine was then applied to other projects of our working group (on *Eurypelma californicum* and *Nautilus pompilius* hemocyanins), and proved to be very effective. The adapted projection matching technique is now well established and the standard method in our group after achieving a median resolution of ∼15 – 10 Å.

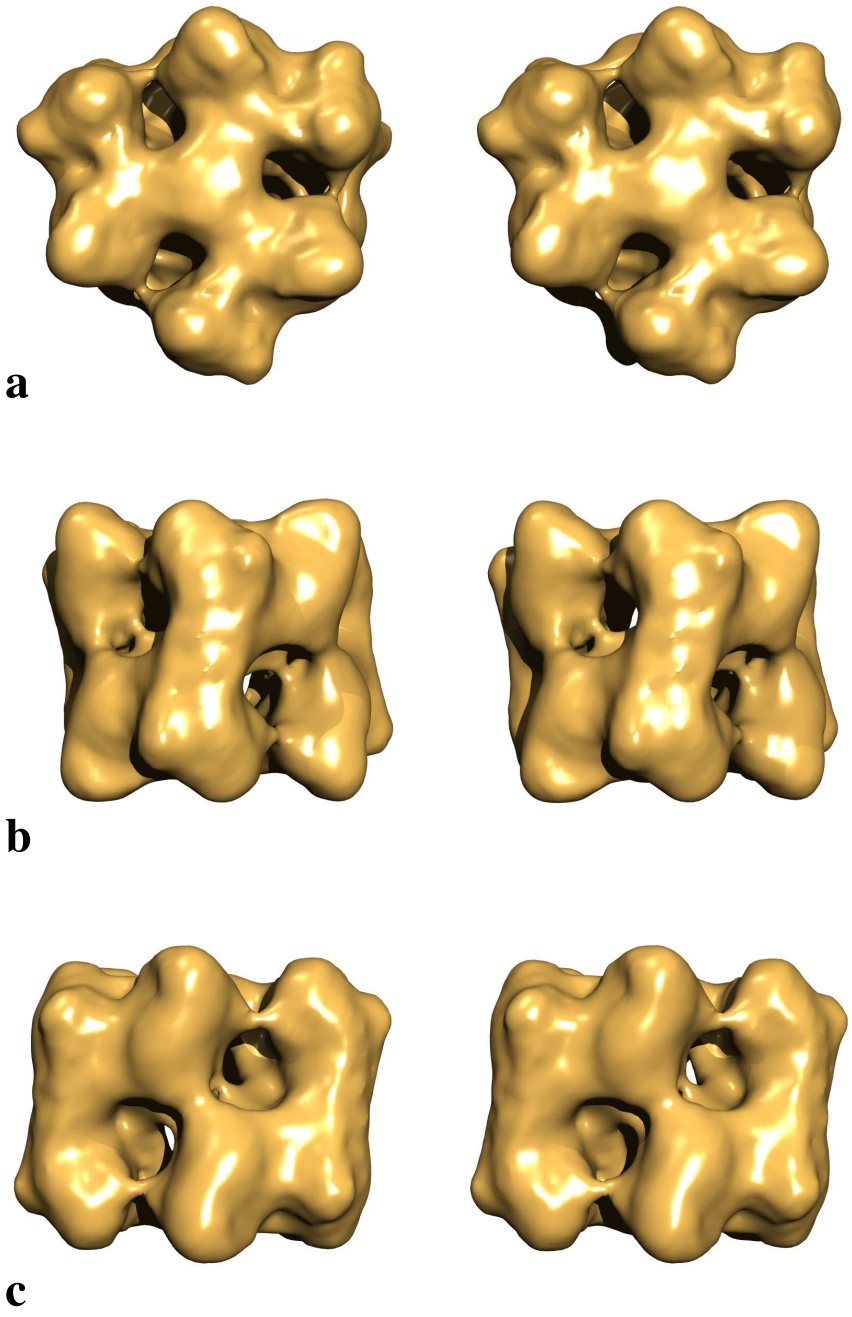
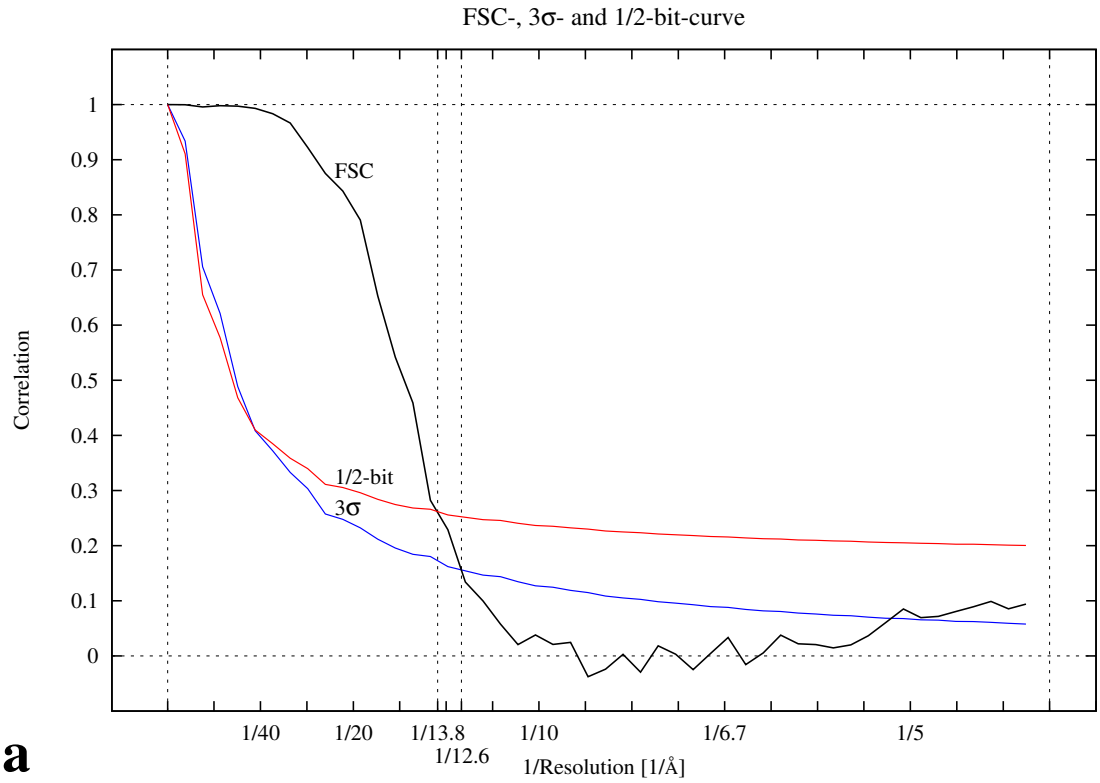
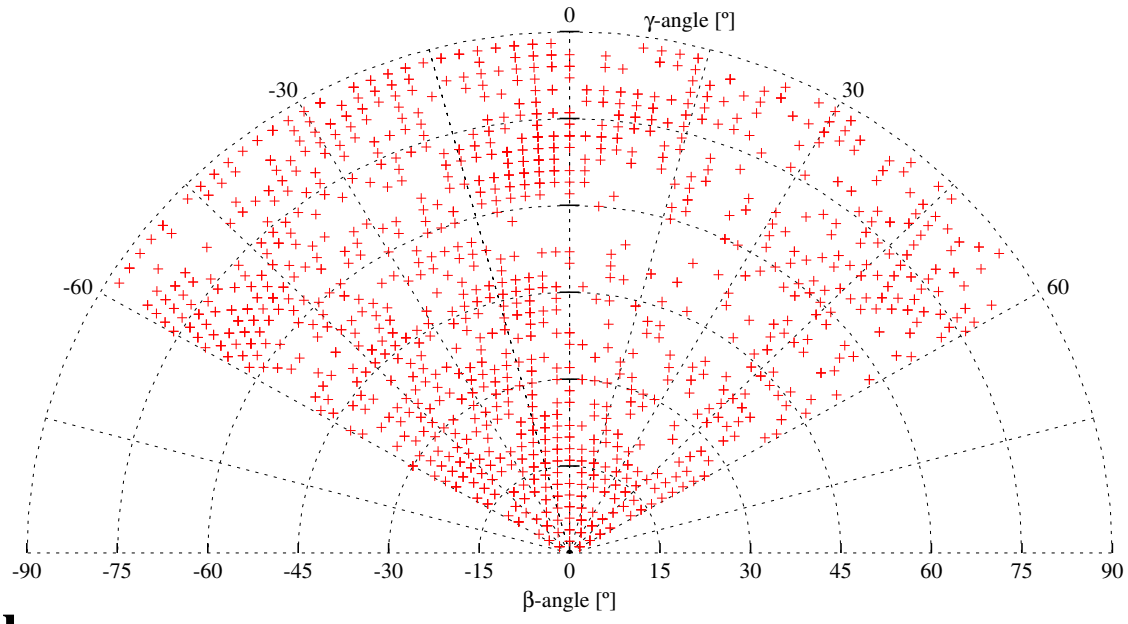


Figure 3.2: Stereo views of the cryo-EM structure of the *Drosophila melanogaster* hexamerin LSP-2 by the technique of angular reconstitution along the main symmetry axes. **a** – *trimer* view (view along the three-fold axis), **b** – *tight dimer* view (view along the first two-fold axis), **c** – *loose dimer* view (view along second two-fold axis).



a



b

Figure 3.3: Determination of the resolution (**a**) and Euler angles distribution (**b**) of the cryo-EM structure of the *Drosophila melanogaster* hexamerin LSP-2 by the technique of angular reconstitution. The distribution of the angles is limited to the asymmetric triangle of the D3 point-group symmetry.

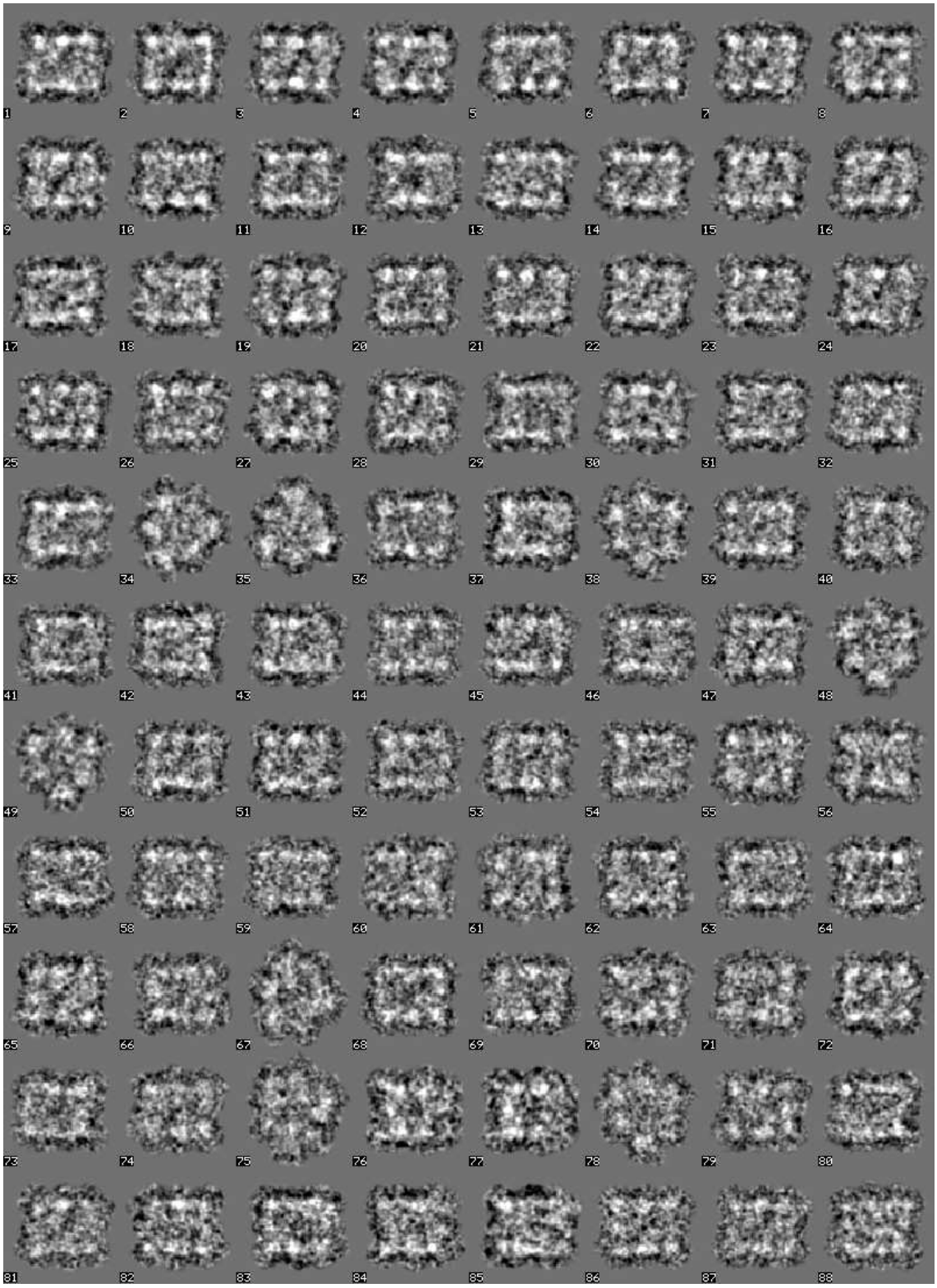


Figure 3.4: The first 88 class averages of the final cryo-EM structure of the *Drosophila melanogaster* hexamerin LSP-2 sorted by ascending cross correlation error. The data set shows a clear predomination of side views (β-angles close to 90°) in the positions with lower cross correlation erros. The first top view appears at position 34 and the first tilted (β-angles between 15° and 75°) view can be seen at position 67.

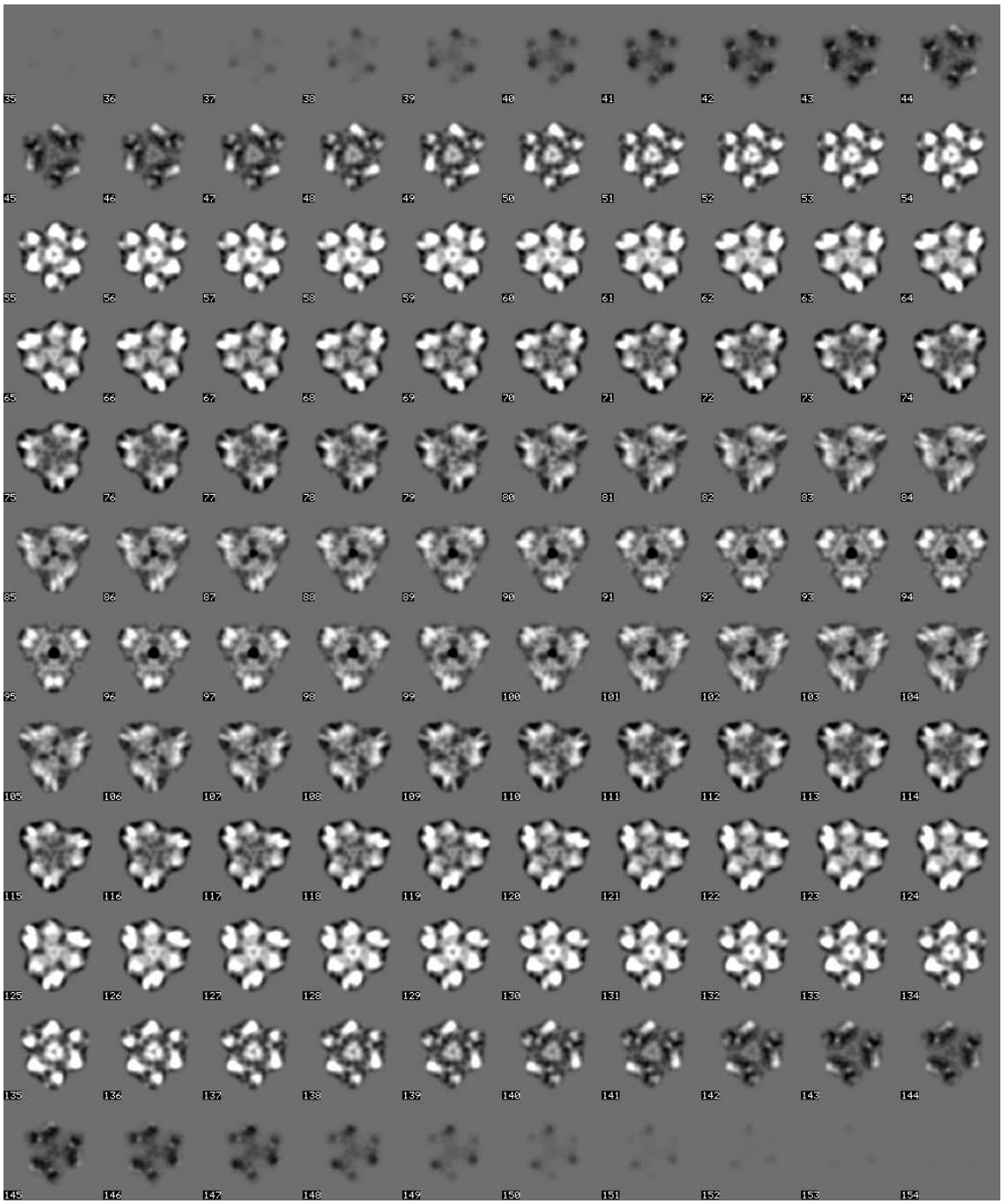


Figure 3.5: Density map of the final cryo-EM structure of the *Drosophila melanogaster* hexamerin LSP-2 by the technique of projection matching. The cryo-EM structure is stacked along the Z axis and every stack has a thickness of ∼1 Å. The central channel along the Z axis is discernible in the stacks 50 – 65 and 120 – 135.

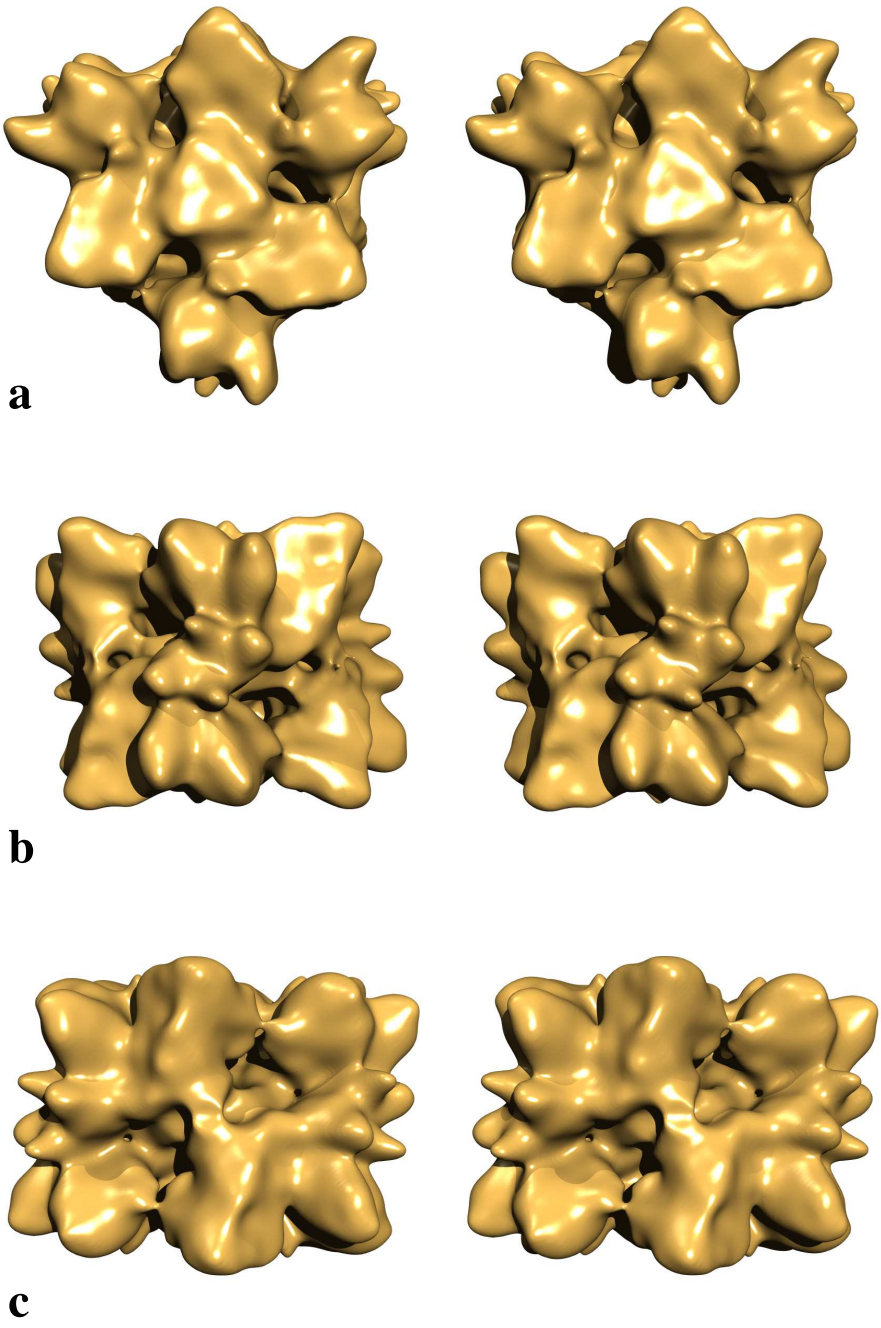
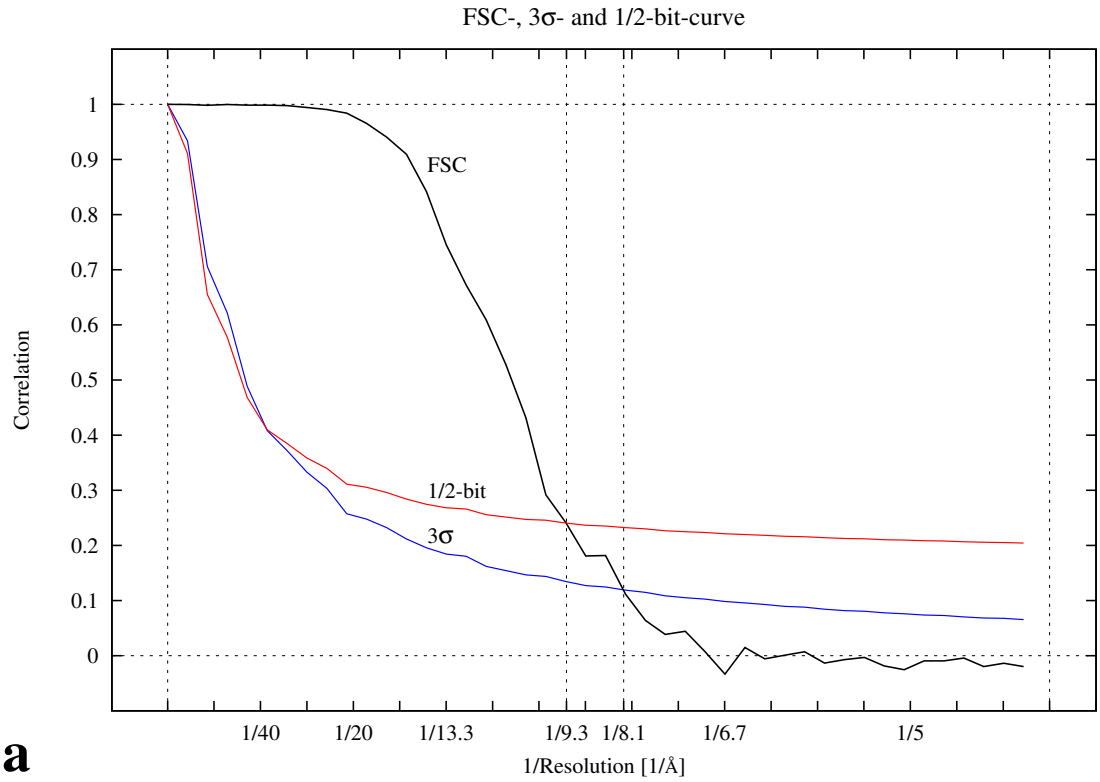
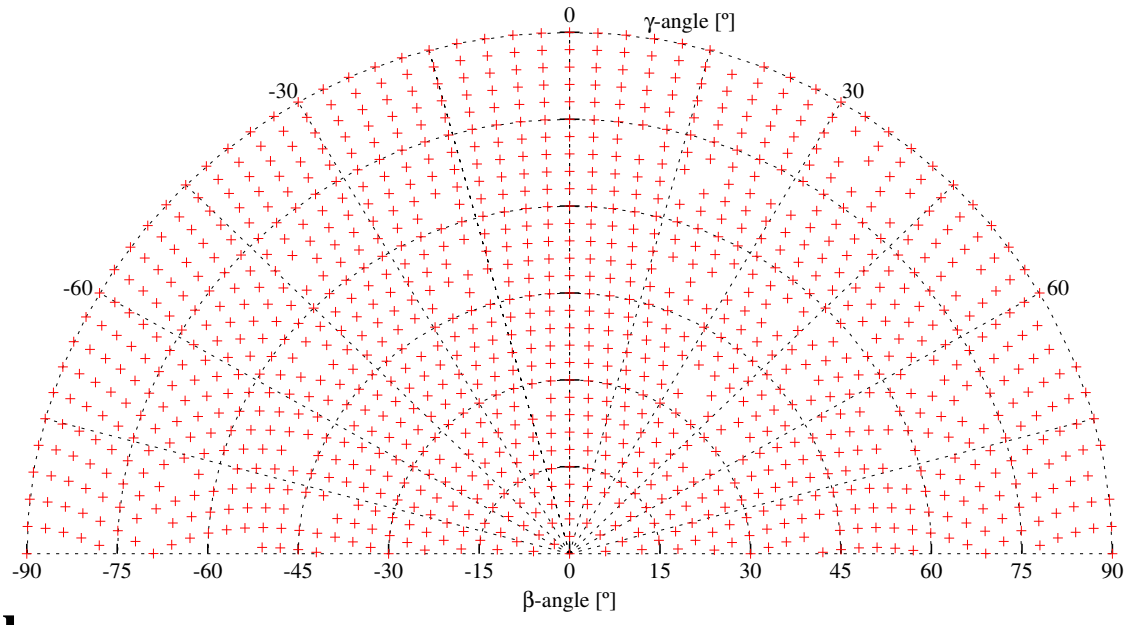


Figure 3.6: Stereo views of the cryo-EM structure of the *Drosophila melanogaster* hexamerin LSP-2 by the technique of projection matching along the main symmetry axes. **a** – *trimer* view (view along the three-fold axis), **b** – *tight dimer* view (view along the first two-fold axis), **c** – *loose dimer* view (view along second two-fold axis).



a



b

Figure 3.7: Determination of the resolution (**a**) and Euler angles distribution (**b**) of the cryo-EM structure of the *Drosophila melanogaster* hexamerin LSP-2 by the technique of projection matching.

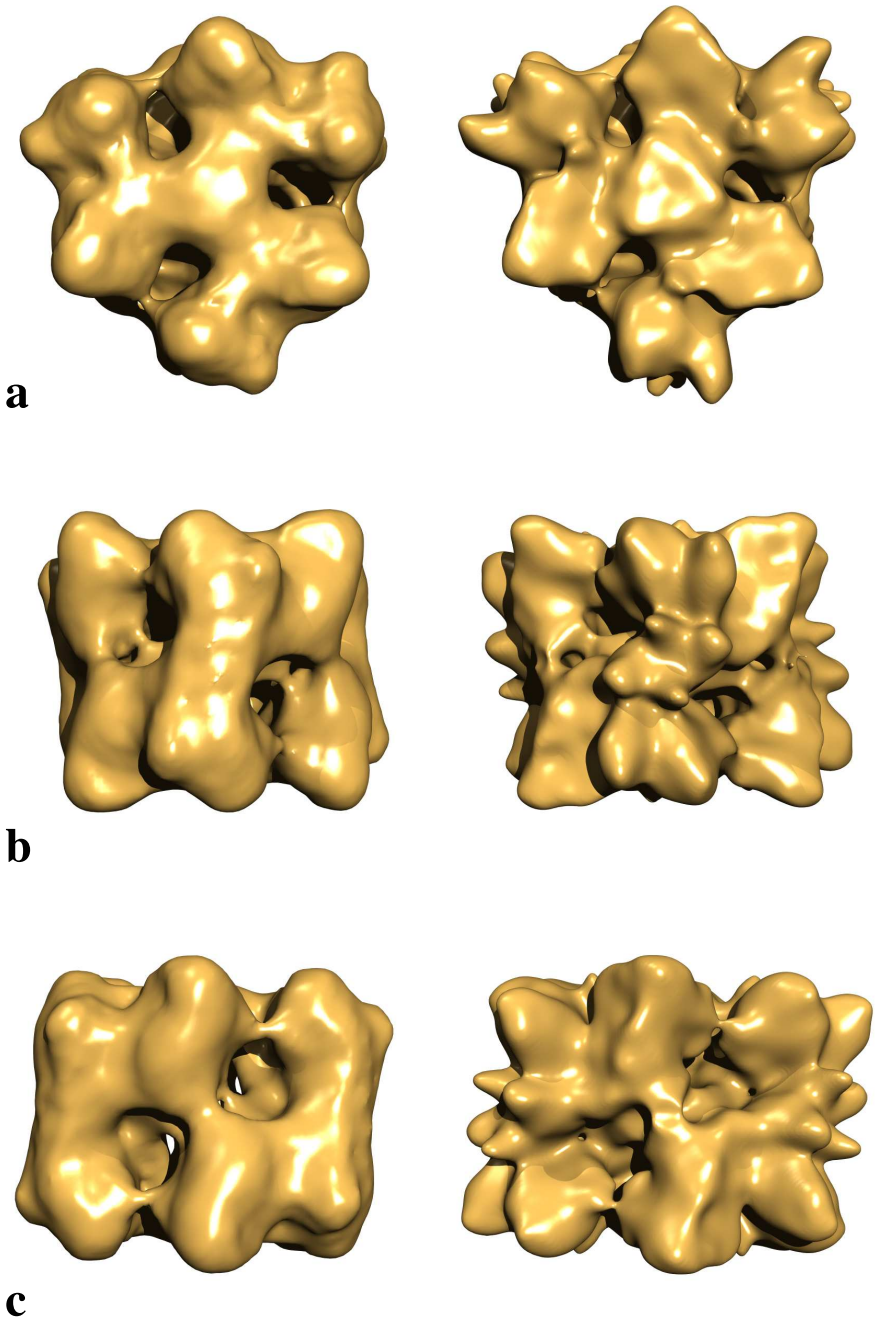


Figure 3.8: Comparison between the results of the technique of angular reconstitution and the adapted technique of projection matching. The left column displays the angular reconstitution result, the right column the result of the adapted projection matching technique. After four rounds of projection matching, resolution improved from 13.8 to 9.3 $Å_{1/2-bit}$ revealing thus many details. **a** – *trimer* view (view along the three-fold axis), **b** – *tight dimer* view (view along the first two-fold axis), **c** – *loose dimer* view (view along second two-fold axis).

3.4 The quaternary structure

The cryo-EM structure of the LSP-2 hexamerin measures 105×135×145 Å (three-fold or *trimer*, first two-fold or *tight dimer*, second two-fold or *loose dimer* axis) and shows six clearly defined subunits. This result is in accordance with the biochemical data presented by [Roberts *et al.*, 1977].

The subunits are bean-shaped and occupy the corners of a trigonal antiprism, yielding a D3 (32) point-group symmetry. In the *trimer* view, the cryo-EM structure appears roughly hexagonal. In the other two views – *tight* and *loose dimer* – the shape is roughly a square and a rectangule, respectively. This result resumes the preliminary data presented by [Meissner *et al.*, 2004] and independently shows that the quaternary structure observed in arthropod 1×6mer hemocyanins ([Volbeda and Hol, 1989a], [Hazes *et al.*, 1993], [Meissner *et al.*, 2003]) is applicable to the *Drosophila* hexamerin LSP-2. Furthermore, the data presented here stresses the assumption that the arthropod hexameric structure is applicable to all insect storage proteins ([Markl and Winter, 1989], [Markl *et al.*, 1992]). The almost exclusively observed cubic shape in the EM data of earlier studies (i.e., [Mousseron-Grall *et al.*, 1997]) can only be explained with interactions of the protein with the support film: the protein adopted preferentially a side view orientation with the three-fold axis parallel to the support film.

The Figures 3.9 to 3.11 show a comparison of surface representations between the cryo-EM structures of the *Drosophila melanogaster* hexamerin LSP-2 (insect) and spiny lobster (*Palinurus elephas*) 1×6mer hemocyanin (crustacean, [Meissner *et al.*, 2003]), and the 3D-volumes of the X-ray structures of *Limulus polyphemus* hemocyanin II (chelicerate, [Hazes *et al.*, 1993]) and *Panulirus interruptus* hemocyanin (crustacean, [Volbeda and Hol, 1989a]). All structures were low-pass filtered to a nominal 10 Å resolution to match the resolution achieved in the cryo-EM structure of the *Drosophila* hexamerin LSP-2.

Even though the overall shape of the cryo-EM structure of the hexamerin LSP-2 is the same as the arthropod hemocyanin's, it shows some structural peculiarities. In the *trimer* view, the cryo-EM structure shows clear protuberances, which I term 'horns', in the region that corresponds to domain 1 (#1). One of the horns is depicted in Figure 3.9 a with an arrow. Figure 3.10 shows the surface representations of the proteins in the *tight dimer* view. Once again, there are protuberances in the hexamerin LSP-2 in the region that corresponds to #1 (marked with arrows in the Figure 3.10 a) that are absent in the hemocyanins. I term these protuberances 'spikes'.

To quantify the correlation between the structures of hexamerin and hemocyanins, 3D-alignments with isotropic magnification allowance were calculated. Table 3.1 summarizes the results of the 3D-alignments and lists the amino acid sequence identities and similarities accessed with the alignment shown in Figure 3.12.

The reader should note the high 3D-cross correlation values between the cryo-EM structures (0.9024) on the one hand and between the 3D-volumes of the X-ray

	DmeHx-LSP-2†		**PelHc**†		**PinHc**∗	
	CC	Sequence id / sim [%]	CC	Sequence id / sim [%]	CC	Sequence id / sim [%]
LpoHcII∗	0.6673	22 / 39	0.6729	30 / 47	0.9821	28 / 46
PinHc∗	0.7073	24 / 44	0.7096	80 / 89		
PelHc†	0.9024	24 / 44				

Table 3.1: Cross correlations between 3D-volumes and amino acid sequence identities/similarities. The 3D-volumes of *Drosophila melanogaster* hexamerin LSP-2 (DmeHx-LSP-2), *Palinurus elephas* hemocyanin (PelHc), *Panulirus interruptus* hemocyanin (PinHc), and *Limulus polyphemus* hemocyanin II (LpoHcII) were 3D aligned and 3D-cross correlation values (CC) calculated. Amino acid sequence identities (id) and similarities (sim) were determined according to the alignment shown in Figure 3.12. Note the high correlation values between the two 3D-recontructions (†) of DmeHx-LSP-2 and PelHc and between the two generated 3D-volumes of the hexameric X-ray structures (∗) of PinHc and LpoHcII. All structures were low-pass filtered to a nominal 10 Å resolution to match the resolution achieved in the cryo-EM structure of the *Drosophila* hexamerin LSP-2.

structures (0.9821) on the other hand. I expected the highest 3D-cross correlation value between the cryo-EM structure of the *Palinurus elephas* hemocyanin and the 3D-volume of the *Panulirus interruptus* hemocyanin, since both are spiny lobsters and the amino acid sequence identity and similarity are the highest in Table 3.1. The calculated value of 0.7096 is lower than the one between the 3D-volumes of the X-ray structures and this leads me to the conclusion that the values cannot be interpreted quantitatively when comparing cryo-EM structures to 3D-volumes of X-ray structures, but only qualitatively. The reason for this discrepancy lies probably in the density variations of the cryo-EM structures, which is a known problem in the docking field ([Chacón and Wriggers, 2002]).

The Table 3.1 shows that a high correlation of quaternary structures can exist, even though the amino acid sequence identity and similarity are rather low. This is clearly seen in the case of the comparison of the cryo-EM structures. Amino acid sequence identity and similarity are low with 24% and 44%, respectively, but the 3D-cross correlation value of 0.9024 is very high.

The alignment in Figure 3.12 revealed nine residues involved in the inter-subunit contacts (as defined in [Volbeda and Hol, 1989a]) that are conserved or isofunctionally exchanged in the three classes, namely $L^{66}(58)$, $D^{68}(60)$, $R^{315}(295)$, $I^{320}(300)$, and $Y^{360}(339)$ in the *trimer* interface, $N^{194}(176)$, $G^{275}(255)$, and $Y^{276}(256)$ in the *tight dimer* interface, and $G^{529}(490)$ in the *loose dimer* interface (numbers in superscript are the positions of the residues according to the alignment in Figure 3.12 and numbers in brackets are the residue numbers according to the X-ray structure of *Panulirus*). The suggestion is that these residues are indispensable for the hexamer formation.

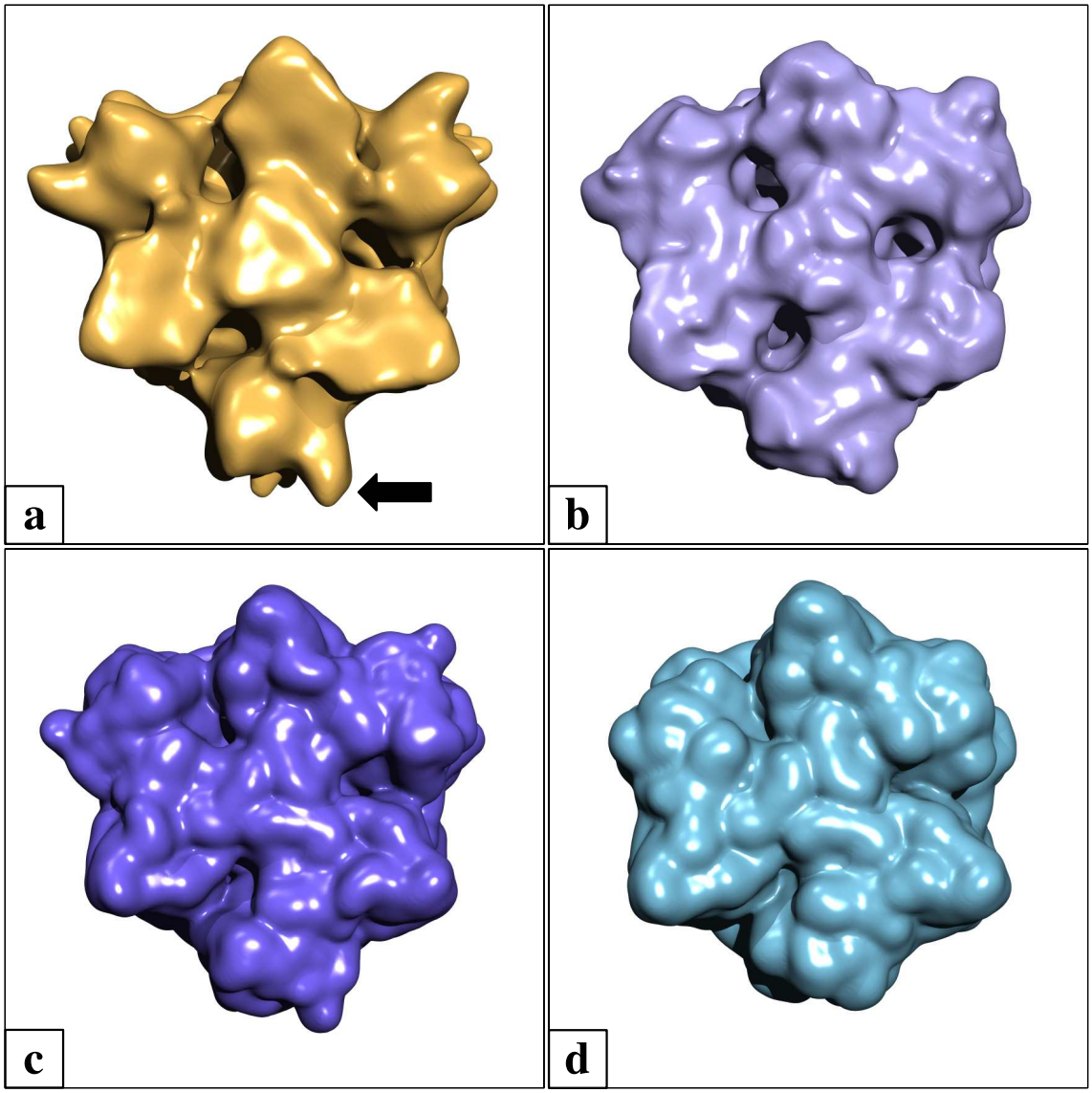


Figure 3.9: Trimer view of different hexameric arthropod proteins. **a** – *Drosophila melanogaster* hexamerin LSP-2 (cryo-EM structure); **b** – *Palinurus elephas* hemocyanin (cryo-EM structure); **c** – *Limulus polyphemus* hemocyanin (3D-volume of X-ray structure); **d** – *Panulirus interruptus* hemocyanin (3D-volume of X-ray structure). All proteins have a roughly hexameric shape in this orientation. The arrow in (a) indicates a protrusion in the region that corresponds to domain 1 and is absent in the hemocyanins. All structures were low-pass filtered to a nominal 10 Å resolution to match the resolution achieved in the cryo-EM structure of the *Drosophila* hexamerin LSP-2.

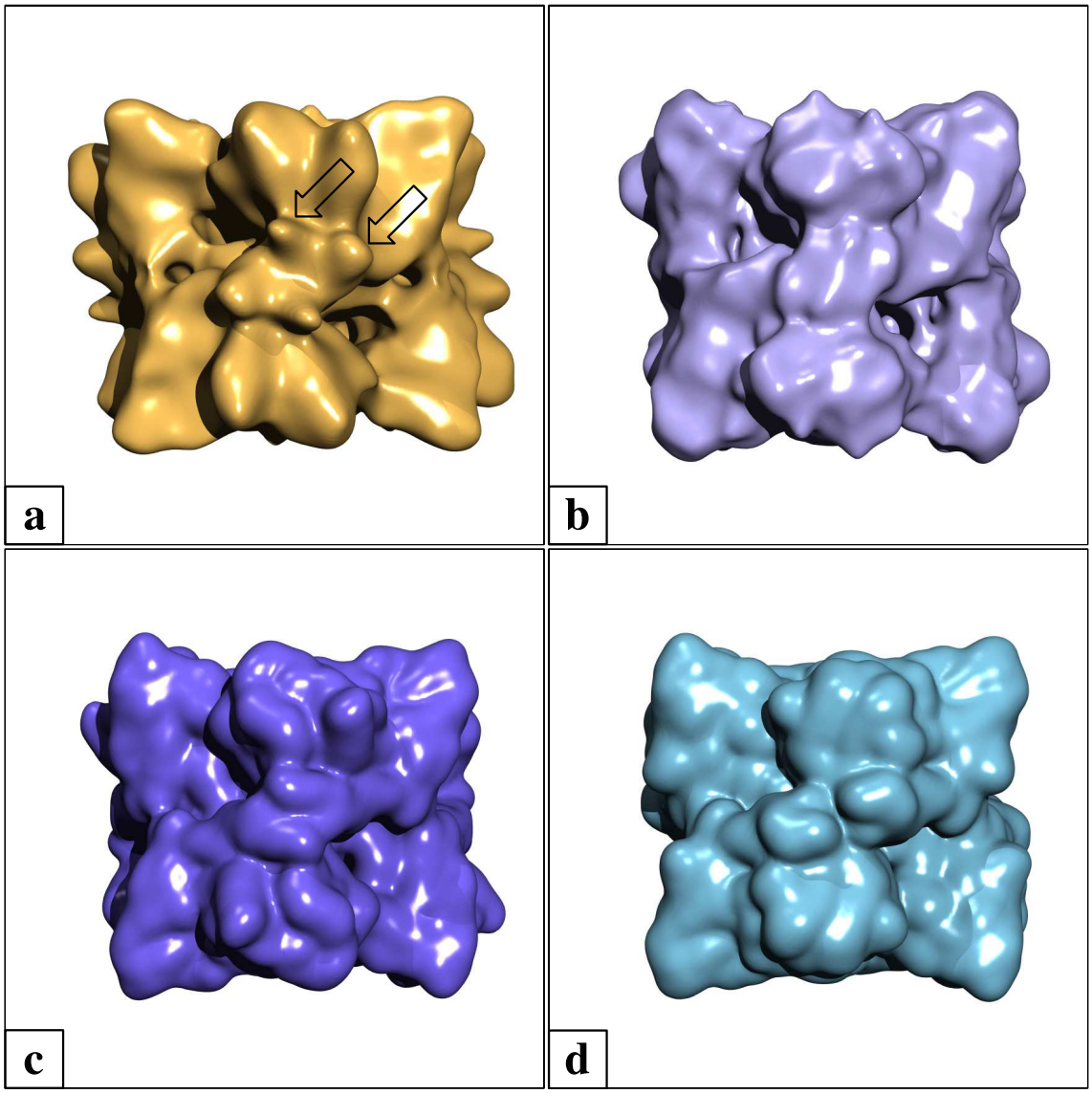


Figure 3.10: Tight dimer view of different hexameric arthropod proteins. **a** – *Drosophila melanogaster* hexamerin LSP-2 (cryo-EM structure); **b** – *Palinurus elephas* hemocyanin (cryo-EM structure); **c** – *Limulus polyphemus* hemocyanin (3D-volume of X-ray structure); **d** – *Panulirus interruptus* hemocyanin (3D-volume of X-ray structure). The shape of the proteins is almost a square in this orientation. The arrows in (a) indicate protrusions in the region that corresponds to domain 1 that are in close contact to the domain 1 of the neighbouring subunit and are absent in the hemocyanins. All structures were low-pass filtered to a nominal 10 Å resolution to match the resolution achieved in the cryo-EM structure of the *Drosophila* hexamerin LSP-2.

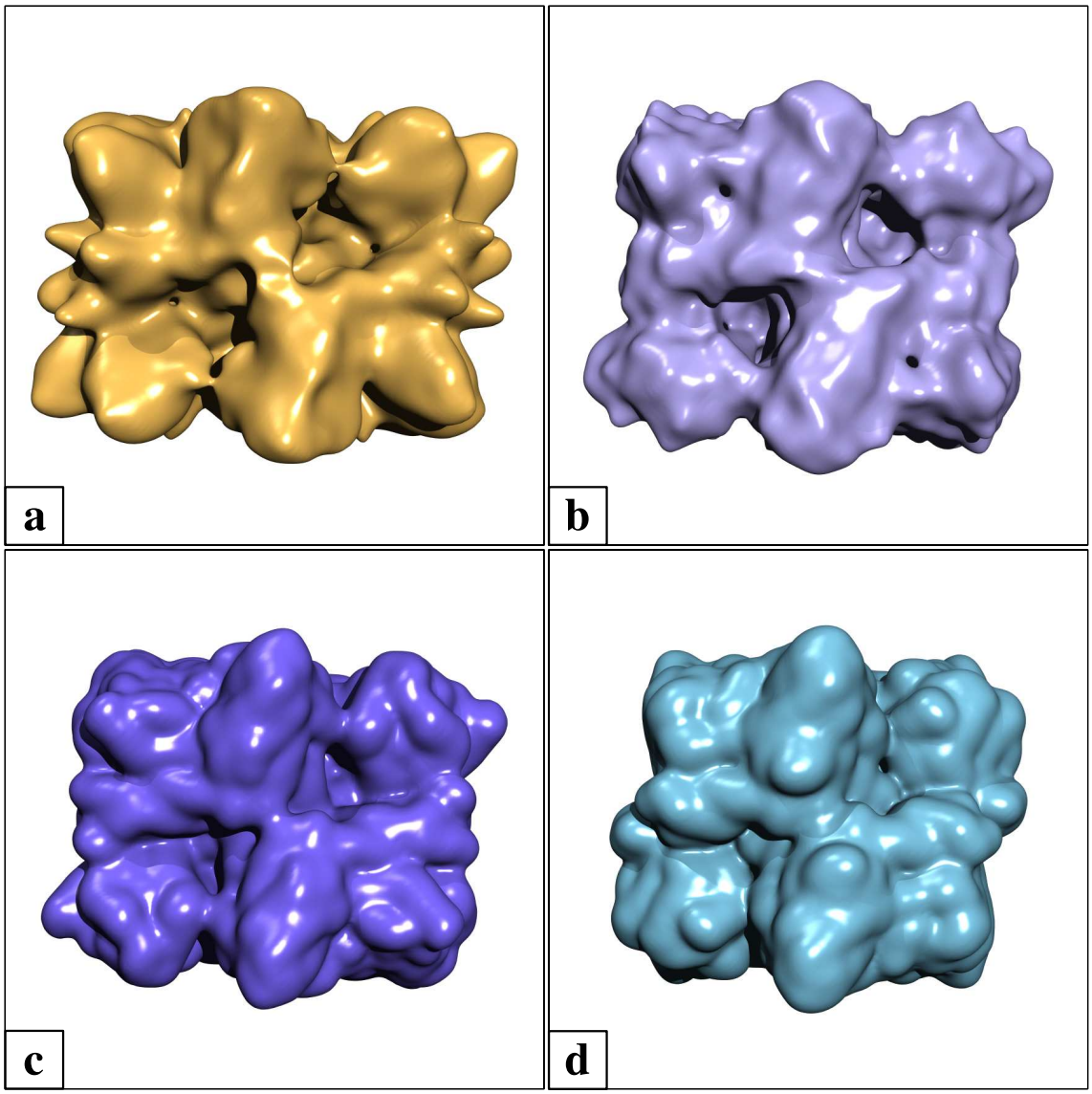


Figure 3.11: Loose dimer view of different hexameric arthropod proteins. **a** – *Drosophila melanogaster* hexamerin LSP-2 (cryo-EM structure); **b** – *Palinurus elephas* hemocyanin (cryo-EM structure); **c** – *Limulus polyphemus* hemocyanin (3D-volume of X-ray structure); **d** – *Panulirus interruptus* hemocyanin (3D-volume of X-ray structure). In this orientation the proteins have a rectangular shape. All structures were low-pass filtered to a nominal 10 Å resolution to match the resolution achieved in the cryo-EM structure of the *Drosophila* hexamerin LSP-2.

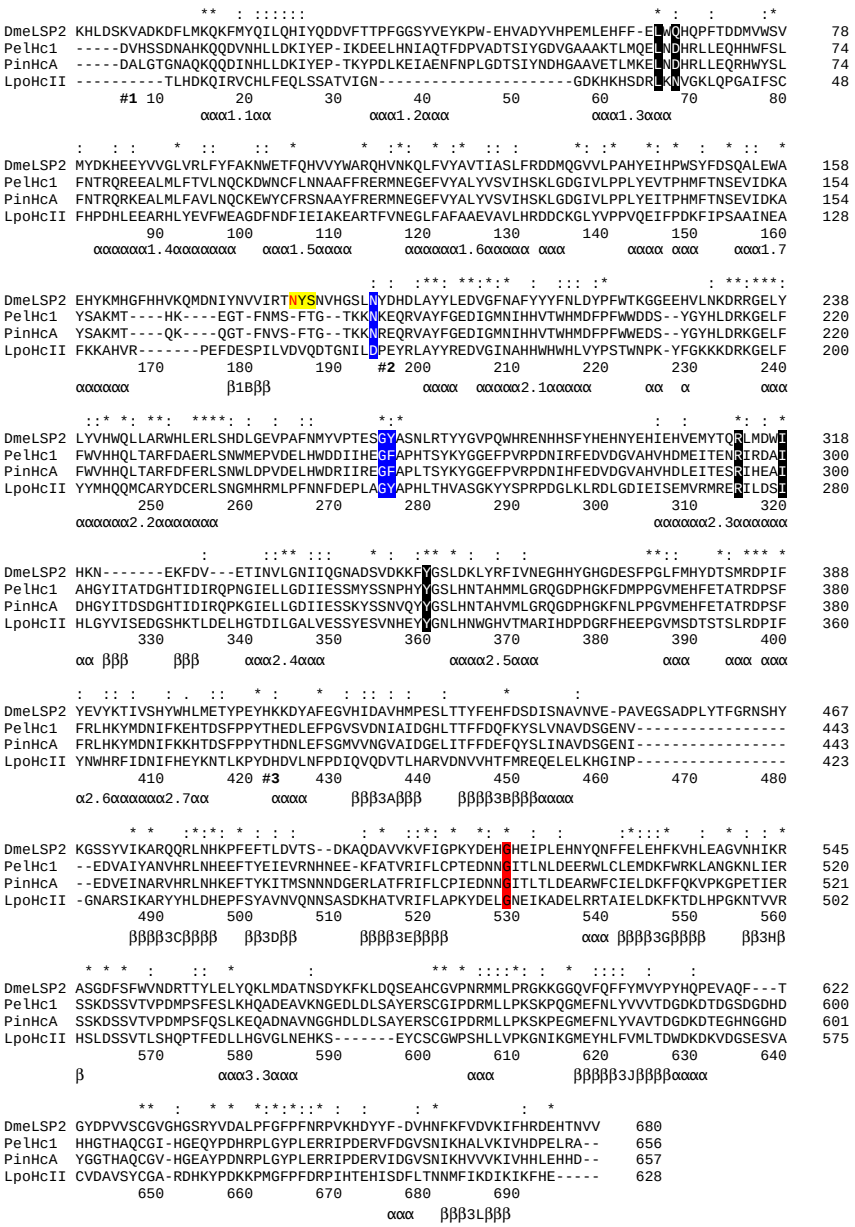


Figure 3.12: Alignment of the amino acid sequences of *Drosophila melanogaster* hexamerin LSP-2 (DmeLSP2), *Palinurus elephas* hemocyanin 1 (PelHc1), *Panulirus interruptus* hemocyanin A (PinHcA), and *Limulus polyphemus* hemocyanin II (LpoHcII). Conserved or isofunctionally exchanged amino acids involved in the inter-subunits contacts (selected according to [Volbeda and Hol, 1989]) are shaded in conformity with their allocation: black – *trimer* interface, blue – *tight dimer* interface, red – *loose dimer* interface. Strictly conserved amino acids are indicated with asterisks (∗), isofunctional amino acids with colons (:). The secondary structure elements as deduced from the X-ray structure of subunit II ([Volbeda and Hol, 1989], [Hazes *et al.*, 1993], [Magnus *et al.*, 1994]) are indicated at the bottom. #1 – domain 1, #2 – domain 2, #3 – domain 3. The Asn-X-Ser/Thr sequon is highlighted in yellow an the predicted *N*-glycosilated asparagine is depicted in red. The alignment was kindly provided by T. Burmester.

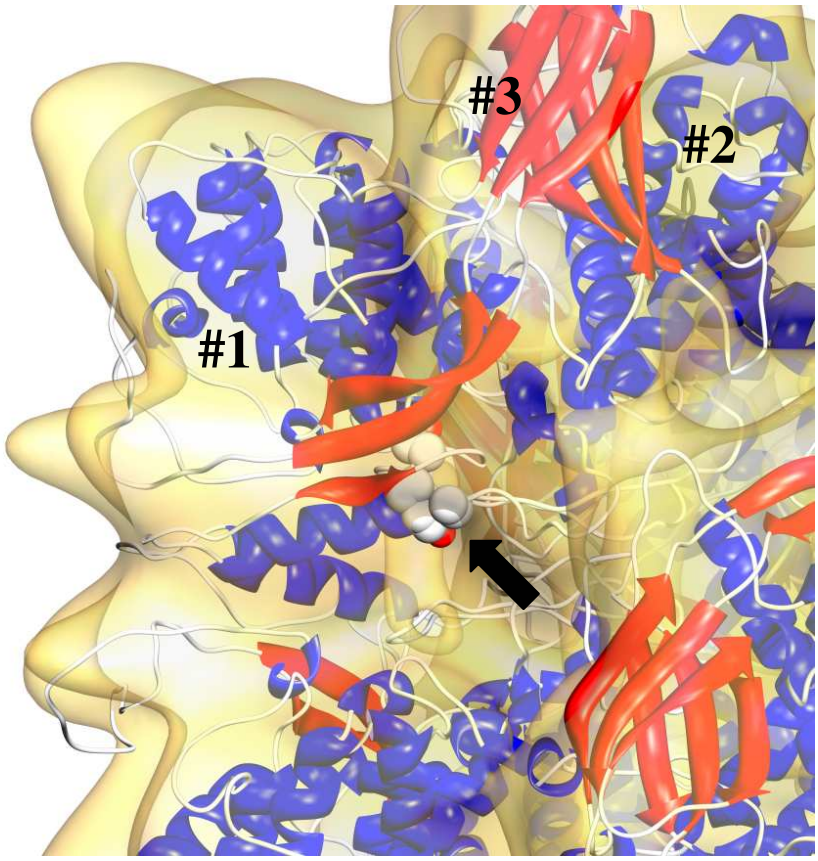


Figure 3.13: Position of the predicted glycosylated N^{185} of the *Drosophila melanogaster* hexamerin LSP-2. The plausibility of the prediction was verified through the molecular model and the cryo-EM structure. The asparagine residue is represented as spheres and indicated by the arrow. The amino acid is situated on the surface of the hexamer in the connection between domain 1 (#1) and domain 2 (#2) and thus most probably glycosylated *in vivo*.

The residue N^{185}(183) is potentially *N*-glycosylated, as predicted by the `NetNGlyc 1.0` Server (artificial neural networks trained on the surrounding sequence context, in an attempt to discriminate between acceptor and non-acceptor sequons). The plausibility of the prediction was verified through molecular modelling (see below), following the study about the hexamerin of the Chinese oak silkworm *Antheraea pernyi* ([Jung *et al.*, 2005]). The asparagine is situated on the surface of the molecule as shown in Figure 3.13 and is most probably glycosylated *in vivo*. This site is in accordance to the first glycosylation site presented by [Mousseron-Grall *et al.*, 1997]. The second glycosylation site presented in latter study was erroneously predicted on an overshooting N-terminal sequence that does not exist.

3.5 Molecular modelling and rigid-body fitting

In the preceding section, two different types of protrusions that are absent in the hemocyanins were adressed: the horns and the spikes. The question is if this structural details are real or artifacts rooting in the manner the hexamerins were isolated and reassociated or in the adapted technique of projection matching. The chosen hexameric model fitted in the final cryo-EM structure answers this question.

As shown in Figure 3.14 a, the horns are situated at the C-terminal end of the amino acid chains (for overviews of the rigid-body fitting see Figure 3.15). The asterisk indicates the first amino acid of the chain in the model an the black oblique arrow points at the volume in the cryo-EM structure that is not occupied by the molecular model. The first nine amino acids (^{1}KHLDSKVAD^{9}) of the LSP-2 sequence were not included in the molecular modelling procedure due to a missing template. Taking into account the local proximity of the C-terminal end of the chain and the unfilled area, I speculate that the missing amino acids of the molecular model are most probably arranged in the horns.

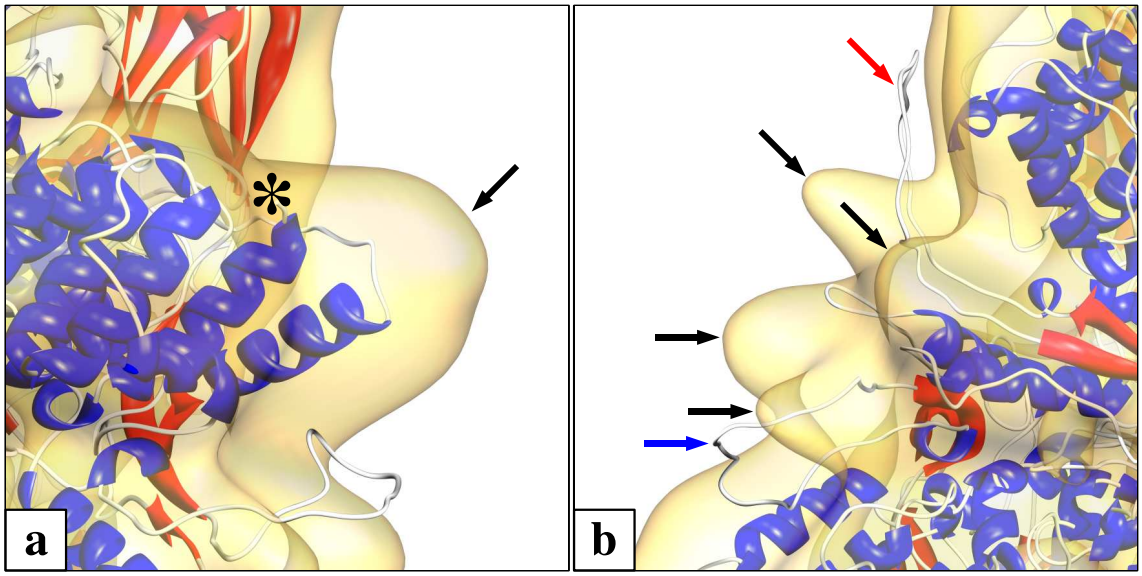


Figure 3.14: Detailed view of the protrusion areas of the *Drosophila melanogaster* hexamerin LSP-2. **a** – Horn area. The asterisk (∗) indicates the C-terminal amino acid of the molecular model and the black arrow points at the area in the 3D-volume that is not occupied; **b** – Spikes area. The coloured arrows mark the not properly modelled insertions of the neighbouring subunits. The oblique and horizontal black arrows point at the areas in the 3D-volume that are not occupied.

The amino acid sequence of the LPS-2 hexamerin has an insertion of 19 amino acids in the β-3B→ β-3C β-hairpin (^{464}VEGSADPLYTFGRNSHYKG^{482}(451–479)). The spikes are located in the area of this insertion (Figure 3.14 b). Accordingly, this unoccupied volume of the cryo-EM structure (oblique and horizontal black arrows) areas are most probably occupied by this insertion. The insertions could not be

properly modelled because of the lacking template and thus stick out of the cryo-EM structure (red and blue arrows). The secondary structure predictions for the insertion yielded a coil region with the possibility of a short β-sheet. This prediction would match the shape of the spikes.

In both cases, there is a high correlation between additional amino acids and unoccupied volumes in the cryo-EM structure. Consequently, I assume that the described features are not artifacts (neither from the biochemical isolation nor from the computational procedure), but real structural details that represent peculiar differences to the hemocyanins.

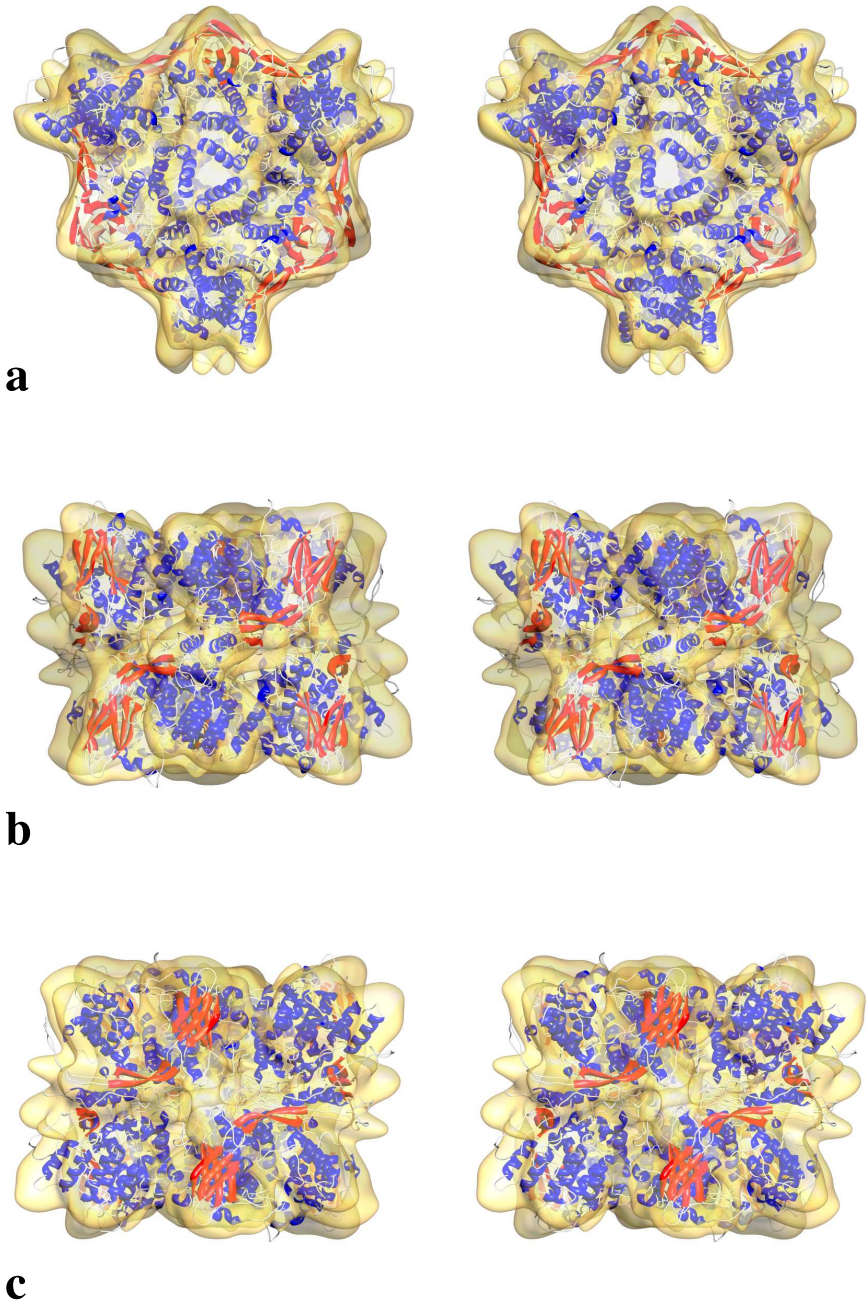


Figure 3.15: Stereo views of the cryo-EM structure with fitted model of the *Drosophila melanogaster* hexamerin LSP-2 along the main symmetry axes. **a** – *trimer* view (view along the three-fold axis), **b** – *tight dimer* view (view along the first two-fold axis), **c** – *loose dimer* view (view along second two-fold axis).

Chapter 4

Results and Discussion of the hemocyanin from *Limulus polyphemus*

4.1 Cryo-EM

A total number of 58 focal pair micrographs were recorded under low-dose conditions (Figure 4.1). Nominal defocus was set between 1.2 and 3.0 μm for the close to focus micrographs and set to 6.0 μm for the far off focus micrographs. With the scanning step size corresponding to 10.6 and 5.9 μm, the digitalised micrographs ended up with a size of 65 and 205 MB, respectively.

In earlier publications, trivial names were given to typical projections were named and this nomenclature is listed with the respective Euler angles in Figure 4.1 a' – e' ([Lamy *et al.*, 1981], [Lamy *et al.*, 1982], [van Heel and Dube, 1994]). The nomenclature is used throughout this chapter.

4.2 Single particle analysis

Single particles were semi-automatically extracted from the micrographs with a high defocus. Single particles from the micrographs with a low defocus were extracted automatically by focal pair alignment. The first data set with a scanning step size of 10.6 μm comprised 8,984 single particles (for both high and low defocus), the second data set with a scanning step of 5.9 μm comprised 3,085 single particles. The estimated real defocus ranged from 1.2 to 4.9 μm.

4.2.1 Angular reconstitution

The analysis started with a reference-free alignment of the 8,984 single particles with a high defocus. The resulting class averages (projections) from multivariate-

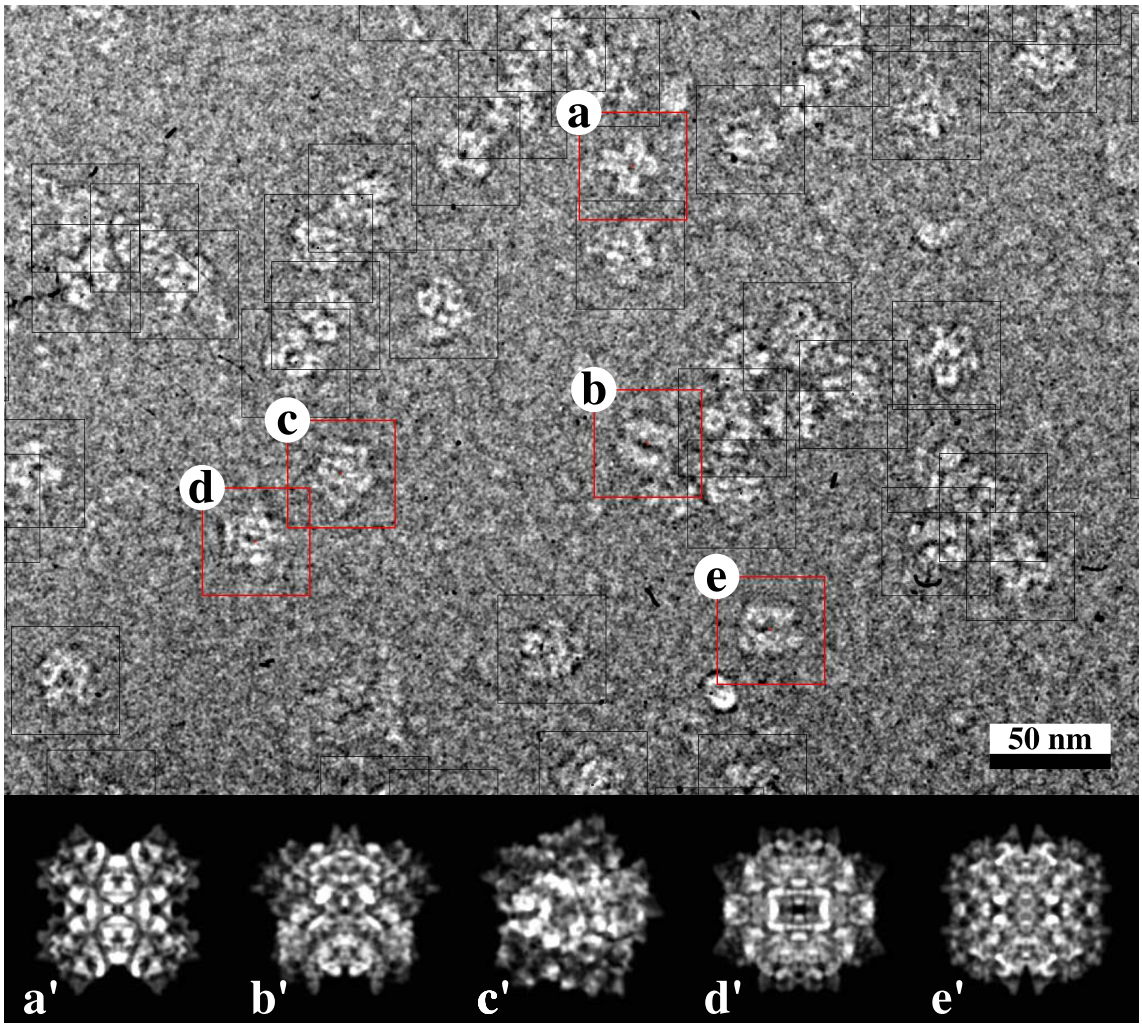


Figure 4.1: Cryo-EM negative and reprojections of the typical views of the *Limulus polyphemus* hemocyanin. The Figure of the cryo-EM negative is a screenshot of the particle selection procedure with the programme `boxer`. The reprojections in the lower row correspond to the highlighted particles in the negative. Typical views are: **a, a'** – Cross view (angles α,β,γ 0°,90°,90°), **b, b'** – Bowtie view (0°,90°,-30°), **c, c'** – Symmetric pentagonal view (0°,45°,-10°), **d, d'** – Ring view (0°,0°,0°), **e, e'** – Lateral view (0°,90°,0°).

statistical analysis (MSA) and hierarchical ascendant classification were used for a first multi-reference alignment (MRA). After two alignment refinements, three projections that represented a top, a tilted and a side view were selected and the Euler angles assigned. These three projections were used as an 'anchor-set' for the remaining data set, and after angular reconstitution a preliminary reconstruction was calculated. Reprojections of the first reconstruction were used for the next round of MRA and MSA. With the first data set with the scanning step of 10.6 μm, a total of 17 rounds of refinements were run. In the first refinement rounds, the number of hemocyanin particles was progressively increased by adding single particles from the

micrographs with a low defocus. After ten refinement rounds, the single particles from the micrographs with a high defocus were eliminated, leaving only images with a low defocus in the data set.

The second data set comprising only images with a low defocus was down-sampled to fit the scanning step size of 10.6 μm of the first data set. Then, both data sets were combined and another five refinement rounds with a total number of 12,069 single particles were run. The whole data set was distributed between 800 class averages (with an average of ∼15 particles per class). The resulting reconstruction from the angular reconstitution technique (Figure 4.2) was calculated with 11,633 single particles averaged in 766 classes and achieved a resolution of 10.8 $Å_{1/2-bit}$ and 8.8 $Å_{3\sigma}$ as shown in Figure 4.3 a. In this study, the resolution of the cryo-EM structure is defined by the 1/2-bit criterion ([van Heel and Schatz, 2005]). The previously applied 3σ criterion yields values that are considered to be rather too optimistic. The threshold was set to correspond to a molecular weight of 3.5 MDa. The distribution of the Euler angles is shown in Figure 4.3 b. The distribution of the projections in the asymmetric triangle is homogeneous with a loose cluster of tilt images with β-angles between -30° and -60° and γ-angles between 0° and 30°. This reconstruction was used for the molecular fitting.

4.2.2 Projection matching

The resulting molecular model of the fitting of the whole 8×6mer was converted into a density map, low-pass filtered to a nominal 5 Å and used as a reference for the final refinement round, applying projection matching. The final cryo-EM structure was calculated with the whole data set (12,069 single particles) in 1140 class averages (∼11 particles per class). The first 88 of these classes (sorted by ascending cross correlation error) are presented in Figure 4.4. There is a clear predominance of tilted projections (β-angles between 15° and 75°) in the first sorted positions. The first cross view (projection along the Y axis) can be seen in position 32, the first lateral view (projection along the X axis) is found in position 58, and the first ring view (projection along the Z axis) appears at position 70. The tomographic layers of the cryo-EM structure are presented in Figure 4.5 and are visualised as stereo surface views in Figure 4.6.

As shown in Figure 4.7 a, the Fourier shell correlation (FSC) curve crosses the $FSC_{1/2-bit}$ noise curve at 9.6 Å and the $FSC_{3\sigma}$ noise curve at 8.1 Å. The distribution of the Euler angles is shown in Figure 4.7 b. The homogeneous distribution of the angles is due to the technique of projection matching. Assuming 0.844 Da/$Å^3$, the threshold was set at 0.01 corresponding to a molecular mass of ∼3.5 MDa (48 subunits of 73,000 Da each). The eight hexamers are clearly defined and exhibit molecular detail. Notably, the 48 individual subunits are discernible (see Figure 4.6). The handedness of the 1×6mer was confirmed by X-ray analysis of *Panulirus interruptus* hemocyanin crystals ([Volbeda and Hol, 1989a]) and this allowed an unambiguous assignment of the correct handedness to the 3D-volume of the 8×6mer.

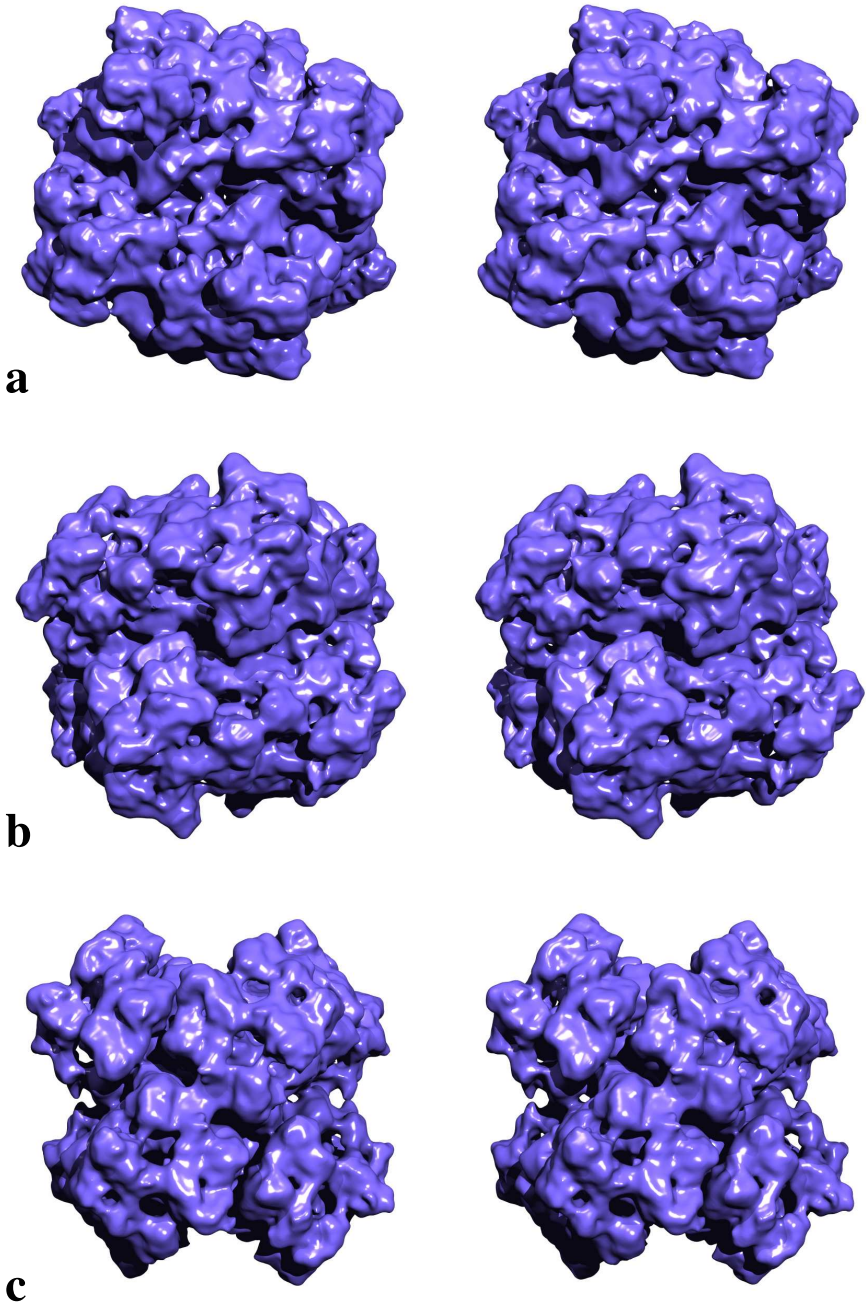
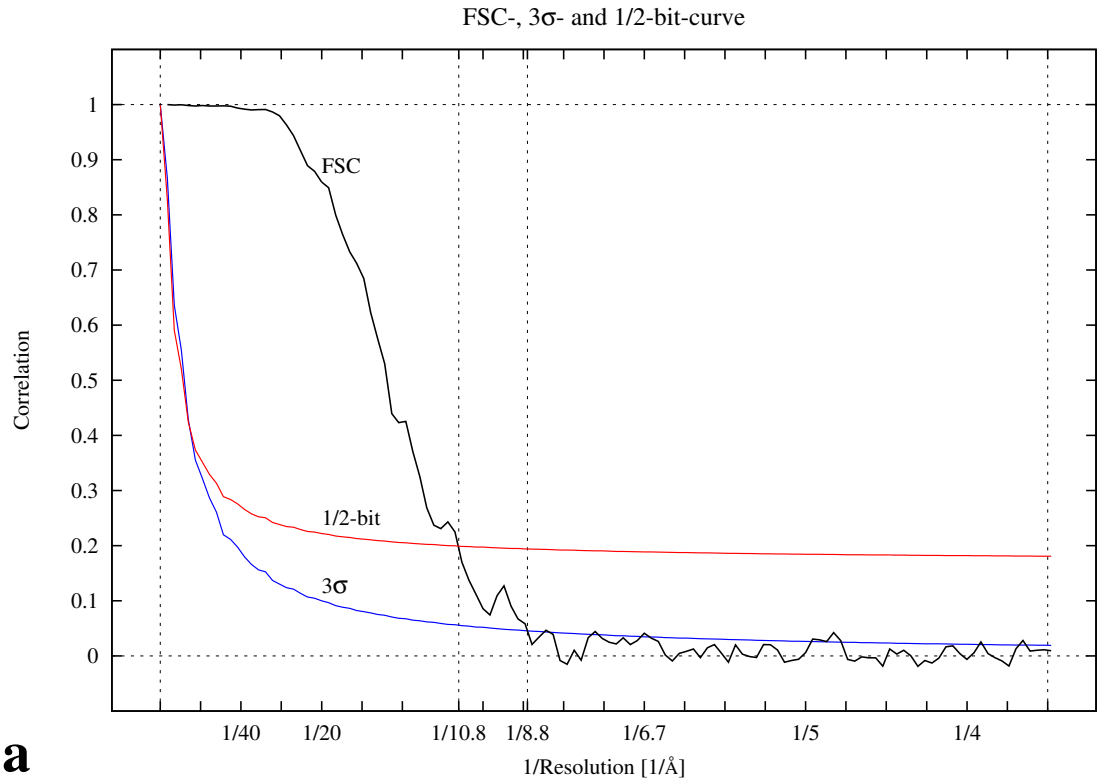
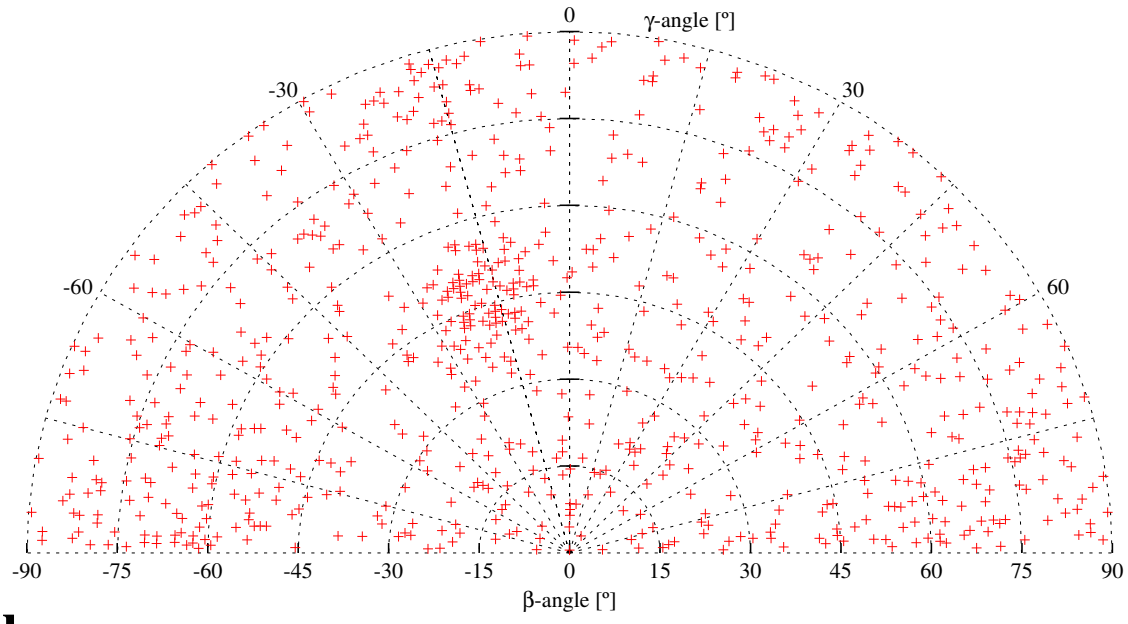


Figure 4.2: Stereo views of the cryo-EM structure of the *Limulus polyphemus* hemocyanin by the technique of angular reconstitution along the main symmetry axes. The subunits are discernible in the clearly defined hexamers of this 10.8 $Å_{1/2-bit}$ resolved cryo-EM structure. **a** – Ring view (view along the Z axis), **b** – Lateral view (view along the X axis), **c** – Cross view (view along the Y axis).



a



b

Figure 4.3: Determination of the resolution (**a**) and Euler angles distribution (**b**) of the cryo-EM structure of the *Limulus polyphemus* hemocyanin by the technique of angular reconstitution. The resolution of the cryo-EM structure of ∼11 Å was achieved with a data set with a nearly homogenous distribution of projections. The homogeneity is disturbed by a loose cluster of projections with β-angles between -30° and -60°.

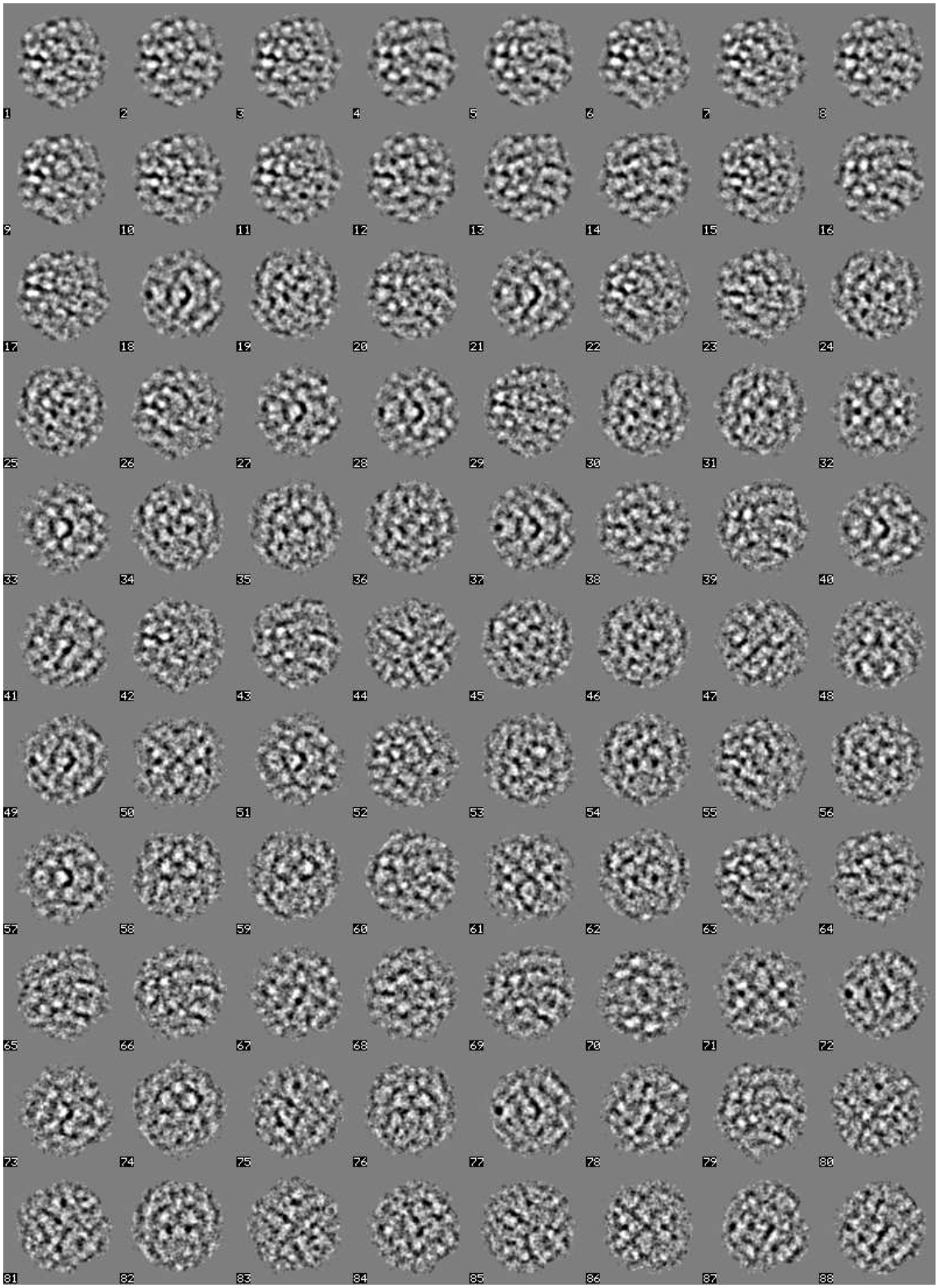


Figure 4.4: The first 88 class averages of the final cryo-EM structure of the *Limulus polyphemus* hemocyanin sorted by ascending cross correlation error. The data set shows a clear predomination of tilted views (β-angles between 15° and 75°) in the positions with lower cross correlation errors. Position 32 – cross view, position 58 – lateral view, position 70 – ring view.

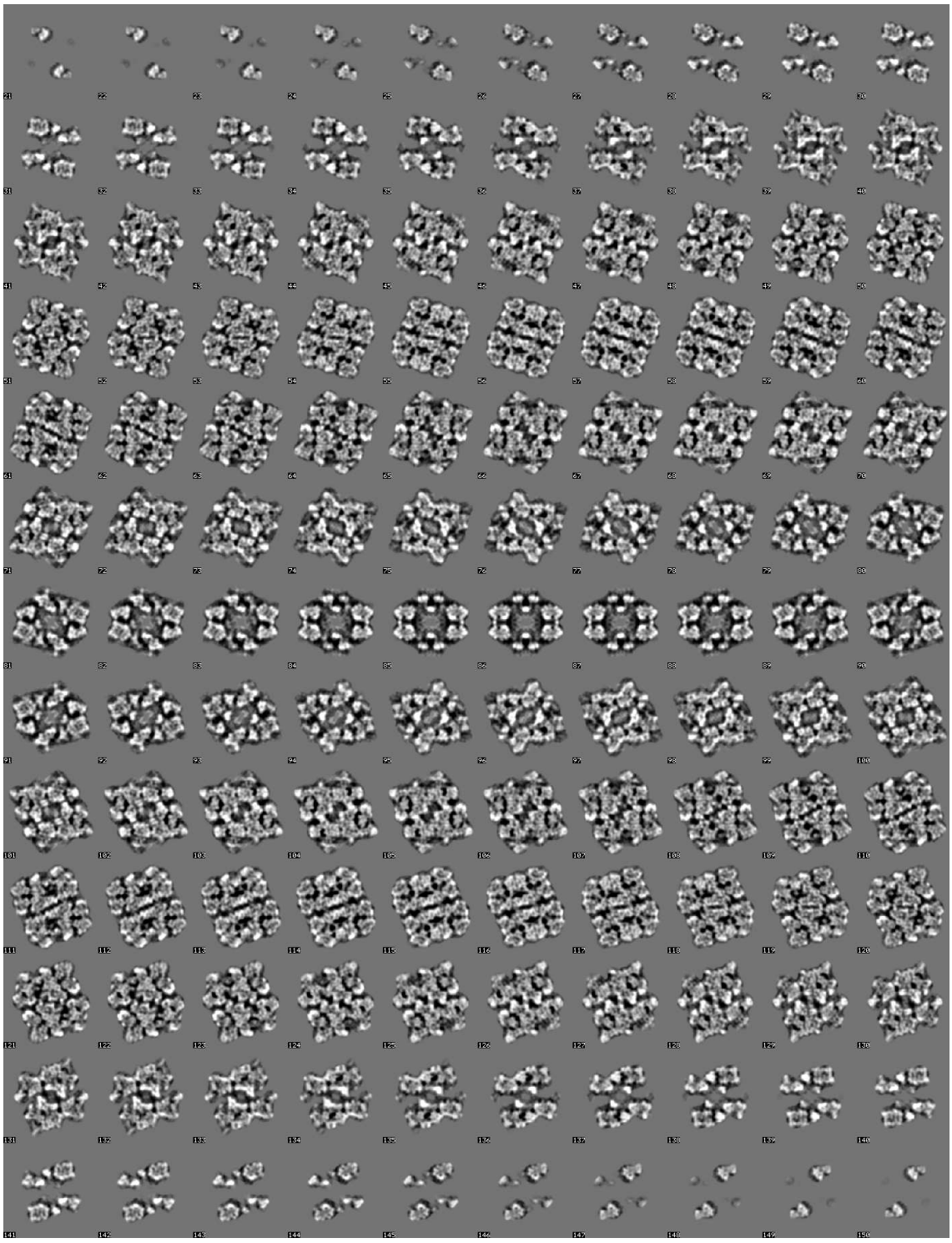


Figure 4.5: Density map of the final cryo-EM structure of the *Limulus polyphemus* hemocyanin by the technique of projection matching. The cryo-EM structure is stacked along the Z axis and every stack has a thickness of ∼2 Å. The staggering angle of 42°(±) is recognizable when comparing stack 95 to 115.

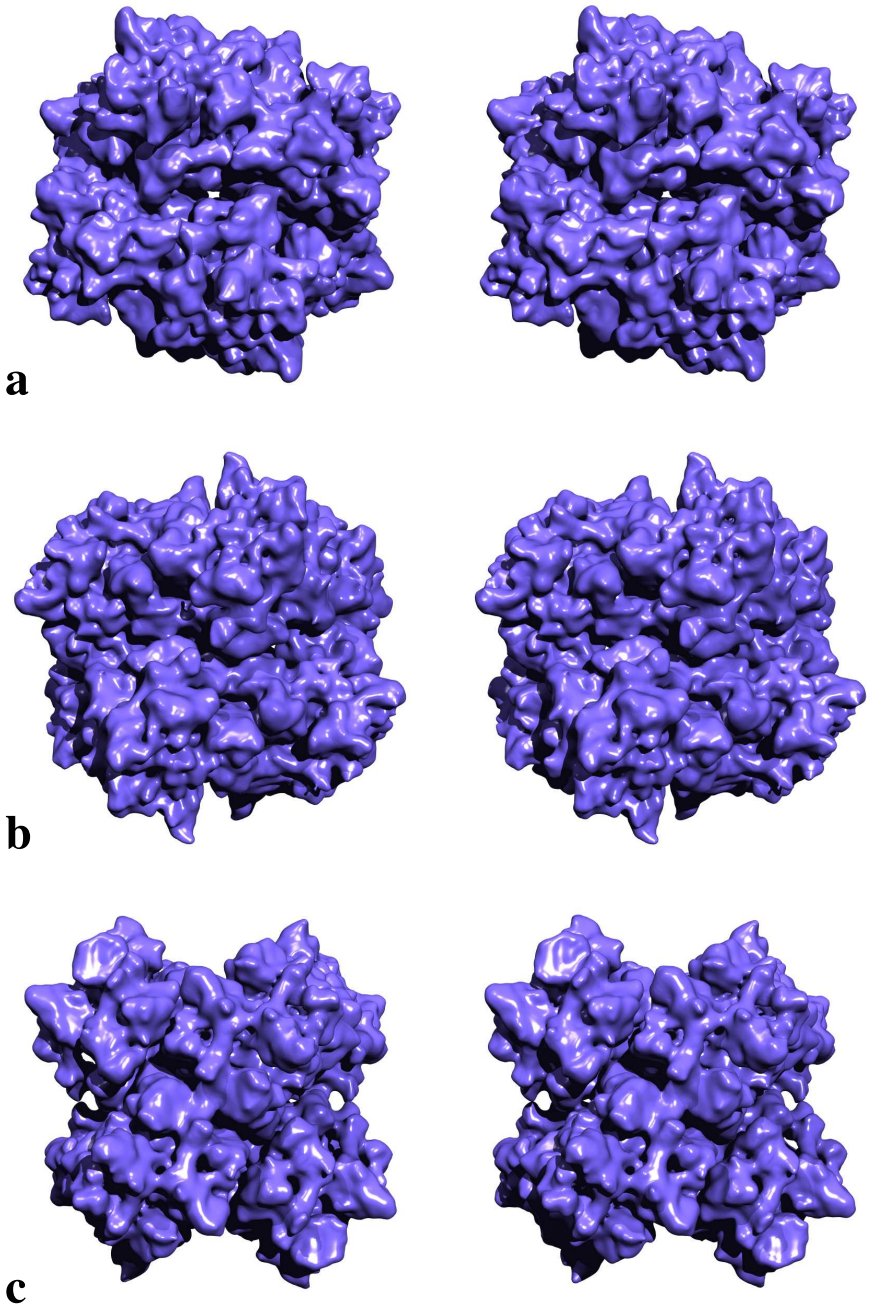
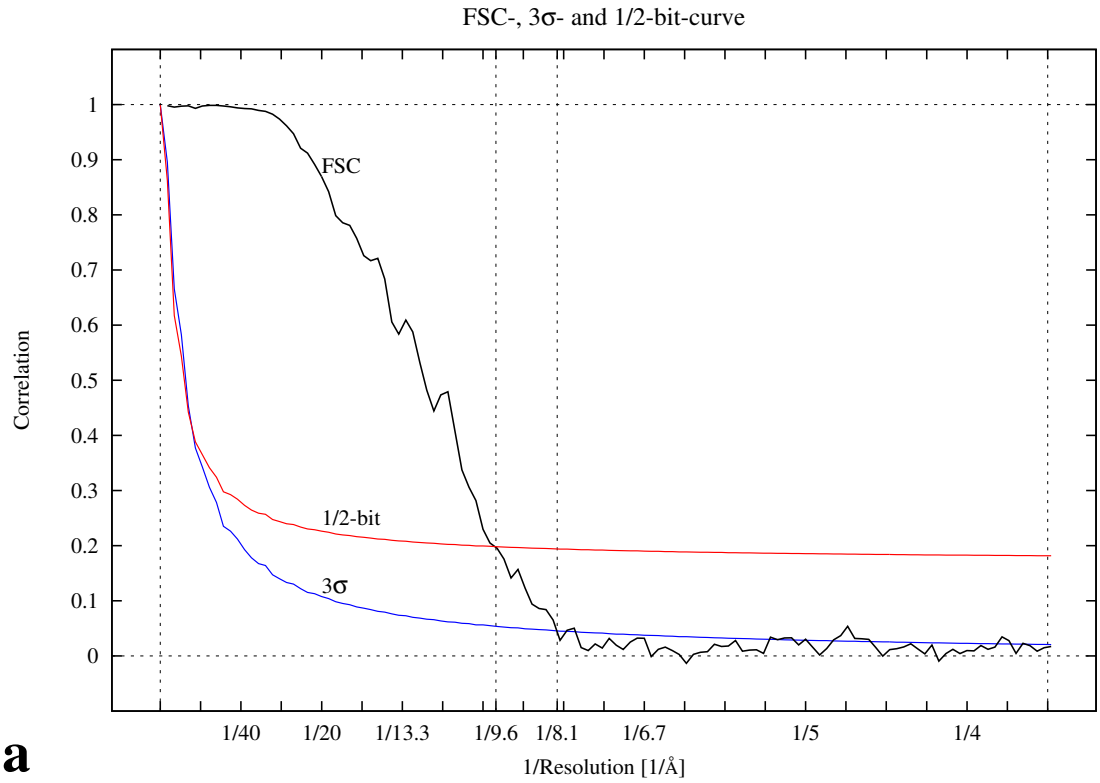
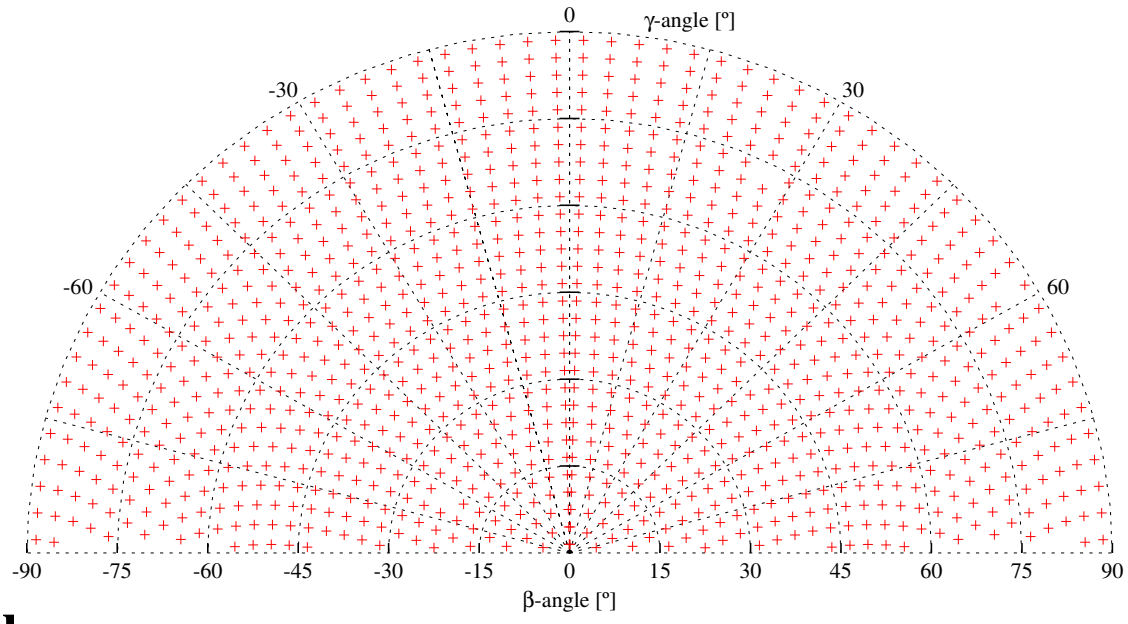


Figure 4.6: Stereo views of the cryo-EM structure of the *Limulus polyphemus* hemocyanin by the technique of projection matching along the main symmetry axes. Compared to the result of the technique of angular reconstitution (see Figure 4.2), this cryo-EM structure shows a more detailed structure with filigree features. **a** – Ring view (view along the Z axis), **b** – Lateral view (view along the X axis), **c** – Cross view (view along the Y axis).



a



b

Figure 4.7: Determination of the resolution (**a**) and Euler angles distribution (**b**) of the cryo-EM structure of the *Limulus polyphemus* hemocyanin by the technique of projection matching. Due to this technique, the resolution could be improved to 9.6 Å and the distribution of the projections got uniform.

4.3 D2 point-group symmetry

The biochemical data presented by [Lamy *et al.*, 1983] poses a question that was already adressed by [van Heel and Dube, 1994]: is the assumption of a D2 point-group symmetry justified for *Limulus* hemocyanin? According to earlier data, three of the edges of the 4×6mer are occupied by subunits type I and the fourth by subunit type IIA ([Lamy *et al.*, 1983]). In such a constellation, the D2 point-group symmetry could be a problem, depending on the resolution achieved in the cryo-EM structure.

Figure 4.8 shows a superposition of the X–ray structure 1NOL of the *Limulus* hemocyanin and the molecular model of the *Eurypelma* hemocyanin subunit type e that was used for the subunit type I/IIA (see below). Both structures match very well apart from the areas where the amino acids are missing in the X–ray structure (highlighted by arrows).

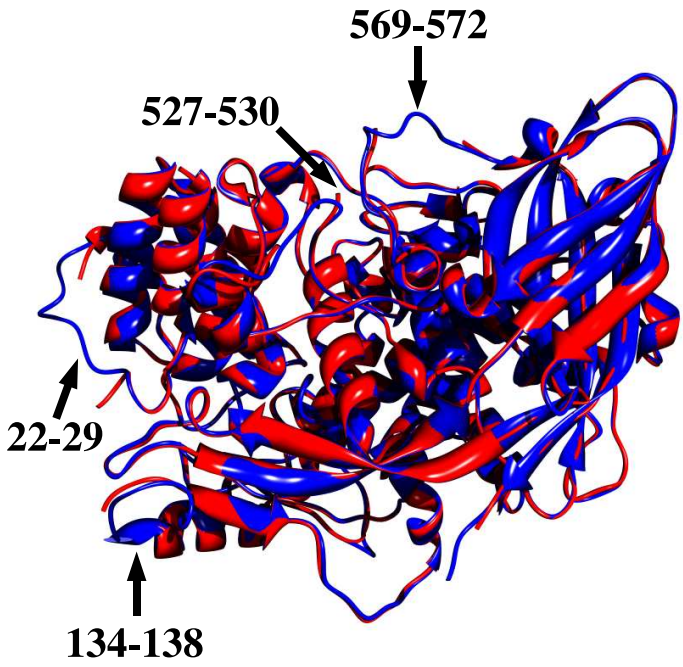


Figure 4.8: Overlay of the X-ray structure 1NOL of the *Limulus* hemocyanin and the molecular model of the *Eurypelma* hemocyanin subunit type e. The X–ray structure is depicted in red and the model in blue. Missing amino acids of the X-ray structure are labelled and the areas marked with arrows. Note the good matching of both structures.

Consequently, at the present resolution of 10 Å, the assumption of the D2 point–group symmetry is not problematic, but this issue should be kept in mind in future studies.

4.4 Quaternary structure of the 8×6mer

Chelicerate 4×6mer and 8×6mer hemocyanins are composed of seven immunologically distinct subunit types that play a specific role in the architecture of the whole molecule. About 25 years ago, several working groups cooperatively elaborated the stoichiometry and topology of these subunit types within the native 4×6mers of *Androctonus* and *Eurypelma*, and the 4×6mer half-structure of *Limulus* hemocyanin, and defined the interspecific subunit correspondencies (for review, see [Markl and Decker, 1992]). For the *Limulus* hemocyanin, the subunit topology was also elaborated at the level of the 8×6mer ([Lamy *et al.*, 1983]). This knowledge allowed me to identify, in the present 3D-volume, the individual subunits types and to label them with their correct designations (Figure 4.9).

In the case of subunit types I, II, IIIA, IIIB, and IV that each occur in each hexamer as a single copy, this assignment was clearly defined. The sixth position in each hexamer is occupied by either subunit V or VI. These two subunits form a central heterodimeric bridge within the basic 2×6mer, and a central tetrameric ring within the 4×6mer (Figure 4.9 a). This leaves two possibilities for their orientation: the one shown in Figure 4.9 a, or the alternative one (with the labels V and VI exchanged). In case of subunit I the picture is complicated because in two hexamers this subunit type is substituted for a structural variant termed subunit IIA ([Lamy *et al.*, 1983]). This subunit is incorporated in Figure 4.9 a–c but is of no relevance for the present study, because within the *Limulus* 8×6mer, subunits I/IIA occupy the outer edges of the eight hexamers and are not involved in any inter-hexamer contact (see Section 4.5). The known subunit topology of the 8×6mer provided a firm basis for molecular fitting (see below).

The architecture of the 8×6mer has already been elaborated by earlier authors ([Bijlholt *et al.*, 1979], [van Heel and Frank, 1981], [Lamy *et al.*, 1982], [Lamy *et al.*, 1983], [van Heel and Dube, 1994], [Taveau *et al.*, 1997]). It consists of two identical 4×6mers assembled face-to-face, with a staggering angle of ∼40° around the two-fold symmetry axis. Each 4×6mer (that correspond structurally to the native 4×6mer hemocyanins of some arachnids such as *Eurypelma californicum* and *Androctonus australis*) is assembled from two identical dodecamers (2×6mers) lying side by side in an antiparallel orientation, but significantly shifted and skewed with respect to each other. Within each 2×6mer, the two hexamers (1×6mers) are rotated ∼105° with respect to each other and differ slightly in their subunit composition (see Subsection 4.4.3). Each hexamer is composed of two staggered layers, each containing three bean-shaped subunits, and the rotation of each layer around the local three-fold axis of the hexamer is 120° (D3 point-group symmetry). This hexamer appears as a hexagon when viewed from the top (along its three-fold axis) and as a square when viewed from the side (see Subsection 4.4.3).

Oligomer	Data	1×6mers in 2×6mer		2×6mers in 4×6mer			4×6mers in 8×6mer	
		Distance (Å)	Rotation (°)	Distance (Å)	Flip-flop shift (Å)	Rocking angle (°)	Distance (Å)	Staggering angle (°)
4×6mer[a,1]	3D (40 Å)	–	105	–	11	14	–	–
4×6mer[b,2]	2D	–	120	–	5	5±1	–	–
4×6mer[b,3]	$SAXS_{oxy}$	105±1.3	105±7	104.4±1.2	18±3	19±4	–	–
4×6mer[b,3]	$SAXS_{deoxy}$	104.5±0.8	131±3	103.5±0.8	4±2	6±3	–	–
8×6mer[c,4]	2D	102±5	90	108±5	17±4	12±3	102±5	32
8×6mer[c,5]	3D (40 Å)	–	103±2	–	12.8±3	12.5±3	108	40
8×6mer[c,6]	3D (10 Å)	100±3	105±5	102±3	14±1	16±2.5	106±3	42±2

Table 4.1: Structural parameters of the 4×6mer and 8×6mer hemocyanins. [a] *Androctunus* hemocyanin; [b] *Eurypelma* hemocyanin; [c] *Limulus* hemocyanin; [1] [Boisset *et al.*, 1995]; [2] [de Haas and van Bruggen, 1994]; [3] [Hartmann and Decker, 2002]; [4] [van Heel and Dube, 1994]; [5] [Taveau *et al.*, 1997]; [6] This study.

Because of the D2 point-group symmetry of the 8×6mer 3D-volume, a repeating 2×6mer is the basic building block of the reconstruction. Some previously published parameters of 4×6mer and 8×6mer hemocyanins are summarized in Table 4.1, together with the values obtained in the present study.

Level	Parameter	Value
2×6mer	Distance (Å)	100±3
	Rotation around 3fold axis (°)	105±5
	Rotation around first 2fold axis (°)	7±2
	Rotation around second 2fold axis (°)	8±2
	Dimensions (Å)	146×230×168
4×6mer	Distance (Å)	102±3
	Flip-flop shift (Å)	14±1
	Rocking angle (°)	16±2.5
	Dimensions (Å)	249×244×168
8×6mer	Distance (Å)	106±3
	Staggering angle (°)	42±2
	Dimensions (Å)	284×271×282

Table 4.2: Summary of the structural parameters of the different oligomeric levels of the *Limulus polyphemus* hemocyanin. The dimensions are measured along the X, Y, and Z axis, respectively.

The most detailed cryo-EM structure of an 8×6mer previously published has a resolution of ∼40 Å ([Taveau *et al.*, 1997]). Compared to this, the present 10 Å cryo-EM structure allows, with a superior accuracy, the assignment of the different positions, shifts and rotations of the two hexamers within the basic 2×6mer, and of the four 2×6mers within the whole molecule. Measurements were performed by two different approaches. First, densities of 1×6mers, 2×6mers, and 4×6mers were isolated from the total density map, and volume alignments were calculated. Second, the generated and fitted models were aligned against each other. The measurements were assessed within the `ResolveRT` software package and yielded the results summarized in Table 4.2.

4.4.1 Topology of the subunits

The identification of the positions of the different subunit types within the cryo-EM structure of the 8×6mer is mainly based on the results of [Lamy *et al.*, 1983]. The different subunit types are present in a different number of copies: six of type I, eight of type II, IIIA, IIIB, and IV, respectively, two of type IIA, four of type V and VI, respectively. For a better tracking of the different subunit types in their positions, a colour code according to Table 4.3 was assigned. The topography of the different subunits is shown in Figure 4.9.

Subunit type	**No. of copies**	**colour**
I	6	red
II	8	green
IIA	2	pink
IIIA	8	yellow
IIIB	8	blue
IV	8	turquoise
V	4	white
VI	4	black

Table 4.3: Table of the colour code for the different subunit types with the specification of the number of copies of each present in the native 8×6mer.

4.4.2 Morphology of the 1×6mer

Although the X-ray structures of a native crustacean hemocyanin hexamer and of a hexamer assembled from subunit II of *Limulus* hemocyanin are available and show the shape and arrangement of the six subunits at high resolution ([Volbeda and Hol, 1989a], [Hazes *et al.*, 1993]), and significantly, yielded a perfect fit with the recent 10 Å cryo-EM structure of a native 1×6mer ([Meissner *et al.*, 2003]), it cannot be *a priori* taken for granted that this arrangement is exactly copied in the various higher-order assemblies. Indeed, a slight but significant deformation was observed by molecular fitting of the subunit II hexamer structure and the present 3D-volume (Figure 4.10).

The seven-stranded anti-parallel β-barrel of #3 is twisted away from the two-fold as shown with the arrow in Figure 4.10 a, while the β-hairpin that embraces #1 is at the exact same position as the β-hairpin form the subunit II hexamer (Figure 4.10 b). The degree of this twisting depends on the subunit position within the whole molecule and the contacts to subunits in neighboured hexamers. Therefore, this is not a reconstruction artefact, but a consequence of the structural interactions between the eight hexamers. For the determination of parameters of the different hierarchy levels of *Limulus* hemocyanin, and interpretations of inter-hexamer interfaces, it is therefore crucial to deal with a density map of the whole 8×6mer (as in

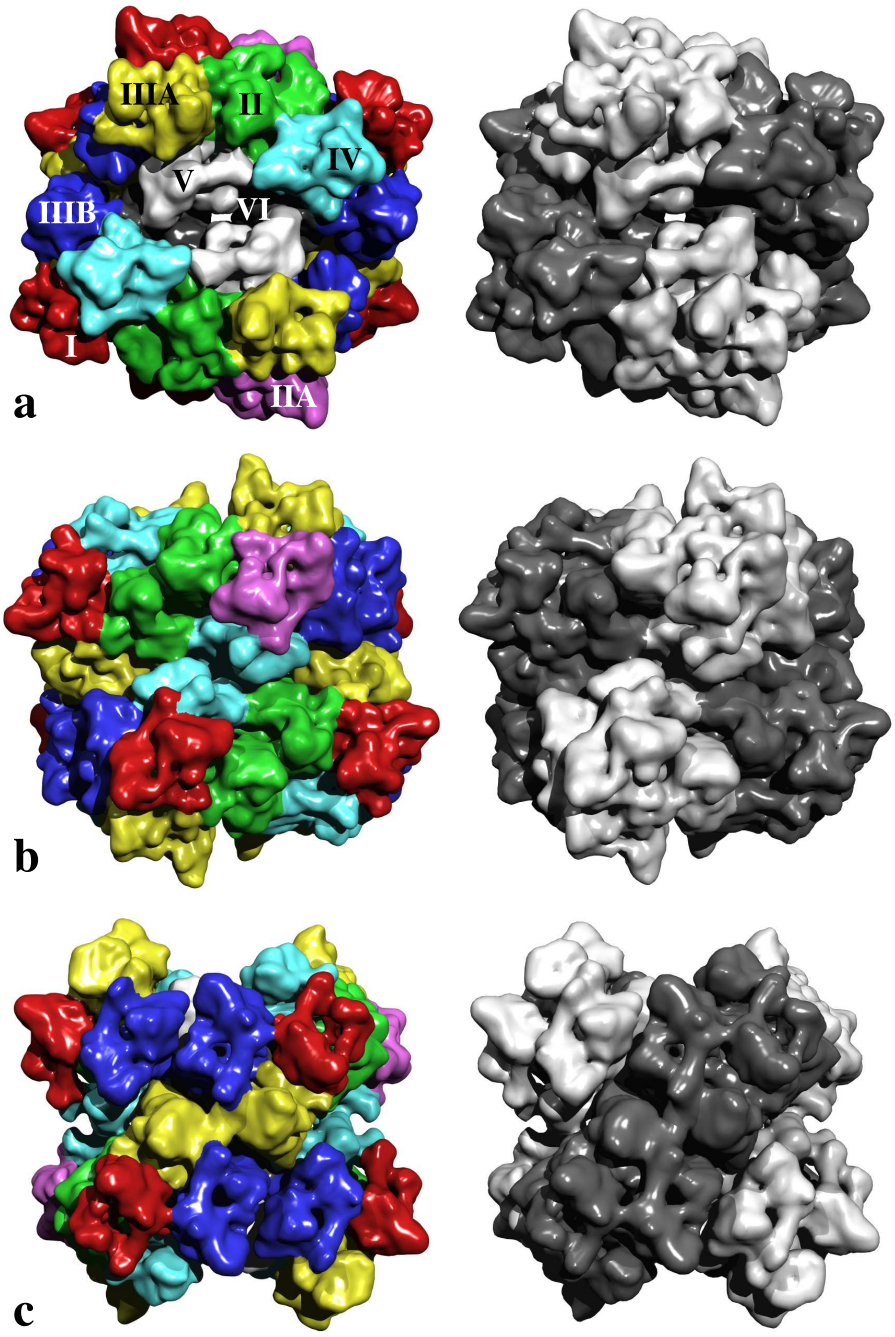


Figure 4.9: Topology of the subunits of the *Limulus polyphemus* hemocyanin. **a** – Ring view (view along the Z axis), **b** – Lateral view (view along the X axis), **c** – Cross view (view along the Y axis). In (a – c) left column, the eight subunit types as biochemically defined ([Brenowitz *et al.*, 1981]), and topologically localized ([Lamy *et al.*, 1983]), are indicated by different colours (red – type I, pink – type IIA, yellow – type IIIA, blue – type IIIB, turquoise – type IV, white – type V, black – type VI). In (a – c) right column, the two hexamers within each 2×6mer are highlighted by light grey and anthracite, respectively. Note that in c, the flop face of the upper 4×6mer is visible

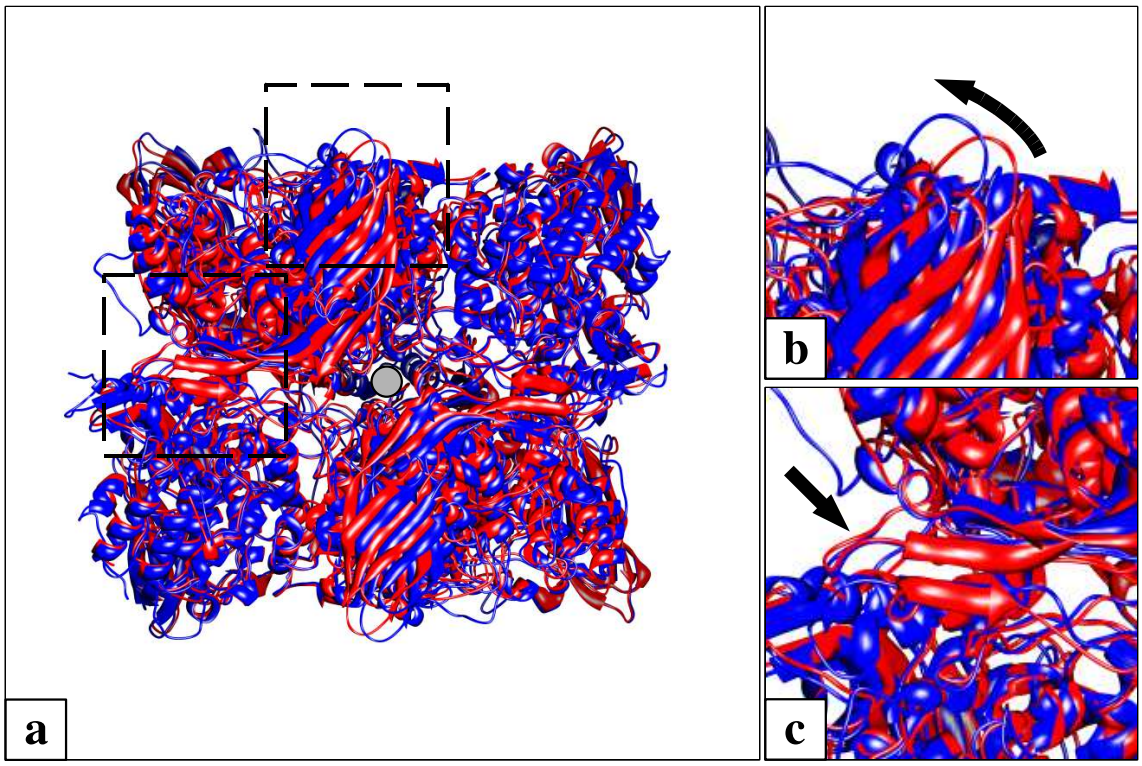


Figure 4.10: Comparison between the X-ray structure of the type II hexamer of the *Limulus polyphemus* hemocyanin (red) and one isolated hexamer of the molecular model (blue). **a** – Overview of the hexamers. The β-sheets of the domains #3 of the molecular model twist away from the two-fold axis (grey disk). The squares delimited with dashed lines indicate the details presented in b and c; **b** – Detailed view of the desplacement of the #3 β-sheets indicated by the curved arrow; **c** – Detailed view of the β-hairpins (arrow). The position of the hairpins is almost identical.

[Taveau *et al.*, 1997] and the present work), instead of a map determined from a combination of sub-structures ([de Haas and van Bruggen, 1994], [Boisset *et al.*, 1995]).

4.4.3 Morphology of the 2×6mer

The 2×6mer measures 146×230×168 Å (X, Y, Z axis) in the 10 Å reconstruction. Within the 2×6mer, the local three-fold axes of the constituent hexamers are 100(±3)Å apart and are not exactly parallel to each other (the upper one is tilted 7° towards the plane of the lower one). The upper hexamer is rotated counter-clockwise 105(±5)° with respect to the lower hexamer around the dodecamer longitudinal axis (Figure 4.11). This gives biochemically isolated *L. polyphemus* 2×6mers, studied in the electron microscope, the appearance of the combination of a hexagon and a square (see for example, Figure 1 bb in [Bijlholt *et al.*, 1979]). Additionally, the upper hexamer is rotated counter-clockwise 8(±2)° around its local three-fold axis, thereby yielding a very close contact to the lower hexamer. This rotation has al-

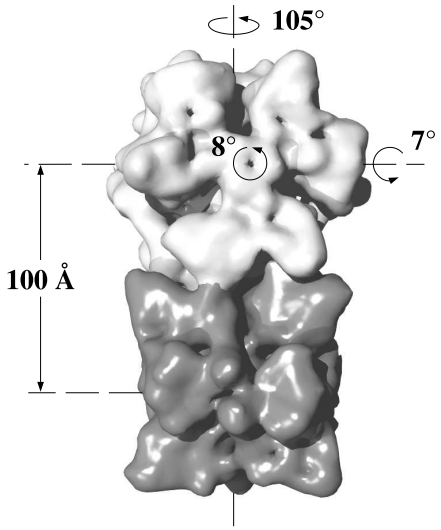


Figure 4.11: 2×6mer, with the distance of the two hexamers, and the rotation of the upper hexamer around its longitudinal axis and around its three-fold axis indicated. In addition, the three-fold axis of the upper hexamer is tilted towards the observer as marked.

ready been described [Taveau *et al.*, 1997]. Note that prior to these rotations, the two hexamers have first to be brought into an antiparallel position with respect to their subunit topology. This is achieved by 180° rotation of one hexamer around its two-fold axis.

4.4.4 Morphology of the 4×6mer

Within the 4×6mer half-molecule, which measures 249×244×168 Å (X, Y, Z axis), the two 2×6mers are associated in an antiparallel arrangement, thereby showing a local two-foldsymmetry (Figure 4.12). Their axial distance is 102(±3)Å. The two dodecamers are shifted 14(±1)Å with respect to each other. This so-called flip-flop-shift is in the range of values previously calculated (see Table 4.1). The flip-flop-shift gives the top view of the 4×6mer the shape of a parallelogram (i.e., with a long and a short diagonal axis, see Figure 4.12 a, a') rather than a square which has first been observed 25 years ago ([van Heel and Frank, 1981]).

As a consequence, two different views are possible, depending on which face of the molecule is exposed to the observer when the 4×6mer rests on a support (Figure 4.12 a, a'). These two orientations have been termed 'flip view' and 'flop view'. In flop view, the lower 2×6mer is on the right side (the long diagonal of the parallelogram runs from lower right to upper left, see Figure 4.12 a). In the flip view, the lower dodecamer is on the left (longer diagonal from lower left to upper right). This should clearly define a 'flip face' and a 'flop face' of the 4×6mer. Unfortunately, Jean Lamy's group and Marin van Heel's group use these terms

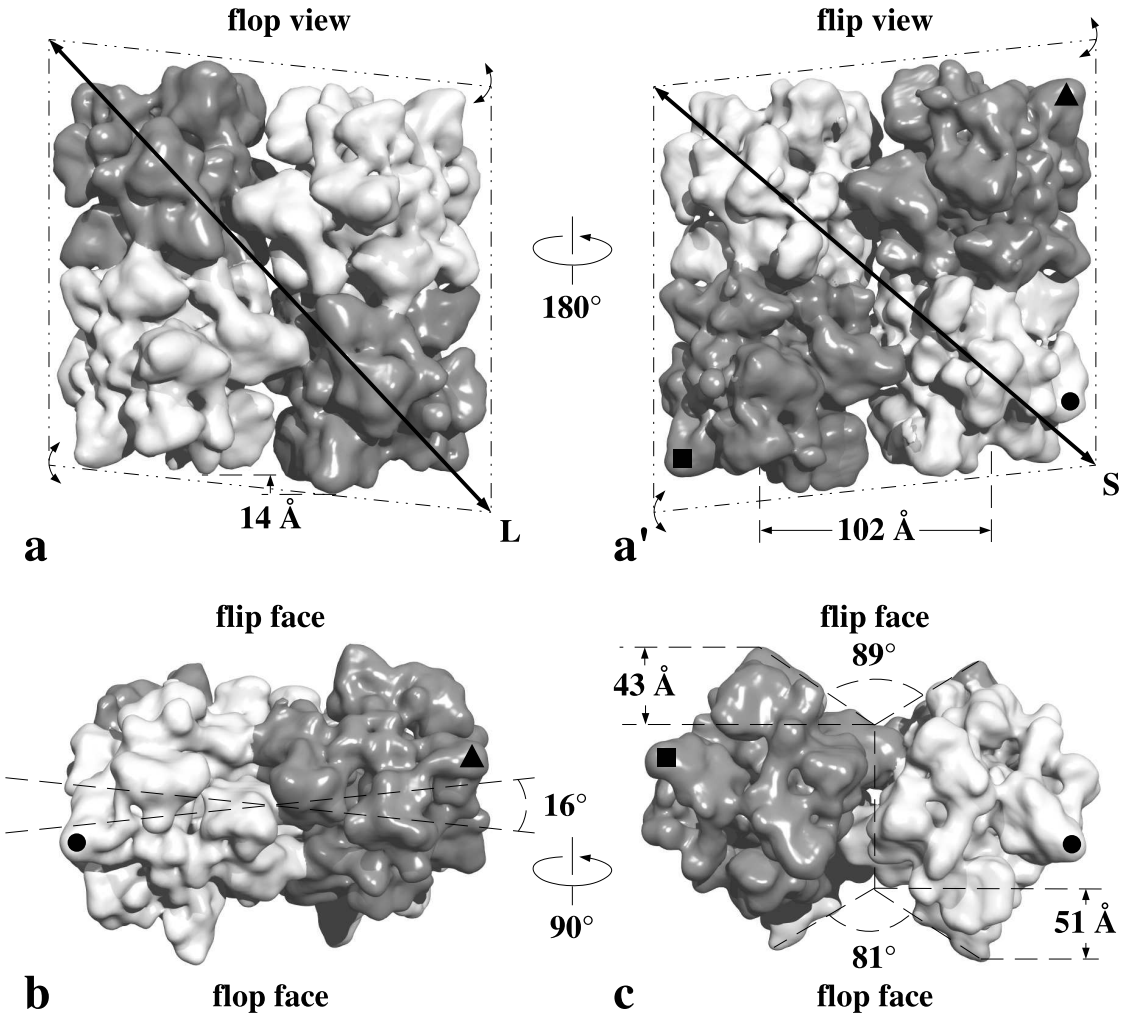


Figure 4.12: Parameters of the 4×6mer. **a** – Flop view of the 4×6mer (exposing its flop face to the reader), with the parallelogram-shape (dashed line), the rocking axis (long diagonal – L), and the flip-flop-shift indicated. **a'** – Flip view of the 4×6mer (exposing its flip face to the reader), with the parallelogram-shape (dashed line), the rocking axis (short diagonal – S), and the distance of the 2×6mers indicated. **b** – 4×6mer viewed along the X axis, with the rocking angle indicated. **c** – 4×6mer viewed along the Y axis (i.e., along the V-shaped clefts between the 2×6mers), with the parameters of the clefts indicated. In a', b and c, three types of small labels mark the corners to facilitate orientation. The terminology of 'flip face' and 'flop face' follows van Heel.

exactly in an opposite manner: According to Lamy, in flip view orientation the flip face of the 4×6mer looks to the support, whereas according to van Heel it looks to the observer! This problem of definition was also addressed by other authors ([Taveau *et al.*, 1997]), and should be kept in mind when comparing the current literature. I decided to give the initial discovery of the flip-flop effect priority and therefore follow here the terminology of van Heel.

The four hexamers are not coplanar, but the longitudinal axes of the 2×6mers are skewed with respect to each other by the so-called rocking angle (Figure 4.12 b). In the 10 Å reconstruction, this rocking angle is 16(±2.5)°. The rocking angle is responsible for the rocking behaviour of the 4×6mer which was first described in 1981 ([van Heel and Frank, 1981]): due to the non-coplanar nature of the 4×6mer, only three of the four hexamers have contact to a plane support. To change between the two stable positions, the molecule has to rock on this support around a diagonal 'rocking axis'. In flop view, rocking occurs around the long diagonal axis whereas in flip view, the molecule rocks around the short diagonal axis (see Figure 4.12 a, a' long and short axes are indicated with L and S, respectively). When a solid molecular model, produced by rapid prototyping is viewed, this rocking behaviour can be easily demonstrated. Dependig on which face is exposed (and thus on which face is lying on the surface), two edges will always rock to and fro around their respective rocking axis. The resulting tilt angle of the 4×6mer (and the 8×6mer) between the two stable positions is the rocking angle.

Between the two assembled 2×6mers, a V-shaped cleft cuts into both faces of the 4×6mer at the median line, but these two clefts differ in their contours: there is a shallow cleft on the flip face and a deep cleft on the flop face (Figure 4.12 c).

4.4.5 Morphology of the 8×6mer

In the 10 Å cryo-EM structure, the 8×6mer measures 284×271×282 Å (X, Y, Z axis). The two 4×6mers have an axial distance of 106(±3)Å (Figure 4.13 a) and are clearly assembled flip-to-flip, thereby exposing their flop faces. Their staggering angle is 42(±2)° (Figure 4.13 b).

Previously, similar shifts and staggering angles were predicted for the flip-to-flip assembly (e.g. [van Heel and Dube, 1994], [Taveau *et al.*, 1997]), keeping in mind that Lamy's group called this arrangement 'flop-to-flop'. The problem of the handedness has to be addressed at each structural level of the molecule (for a detailed discussion of this issue, see [van Heel and Dube, 1994]). The decision between the two possible enantiomers of the 1×6mer was solved by X-ray analysis of the *Panulirus interruptus* hemocyanin crystals ([Volbeda and Hol, 1989a]).

At the next hierarchy level, for creating the typical shape of the 2×6mer seen in the electron microscope (combination of a hexagon and a square), two different ways of rotating one hexamer relative to the other are possible, and for the final two levels the number of possible enantiomers multiplies. In a thorough assessment of all available data, up to the level of the 4×6mer, a single enantiomer was selected as the correct one ([van Heel and Dube, 1994]). This enantiomer was also selected by other authors ([Taveau *et al.*, 1997]) and is definitely confirmed by the present study: only the enantiomers shown in this paper fit the published X-ray structure, whereas the respective mirror images simply do not fit.

In their pioneering work, van Heel & Dube ([van Heel and Dube, 1994]) could not unambiguously decide whether at the 8×6mer level a flip-to-flip or a flop-to-flop

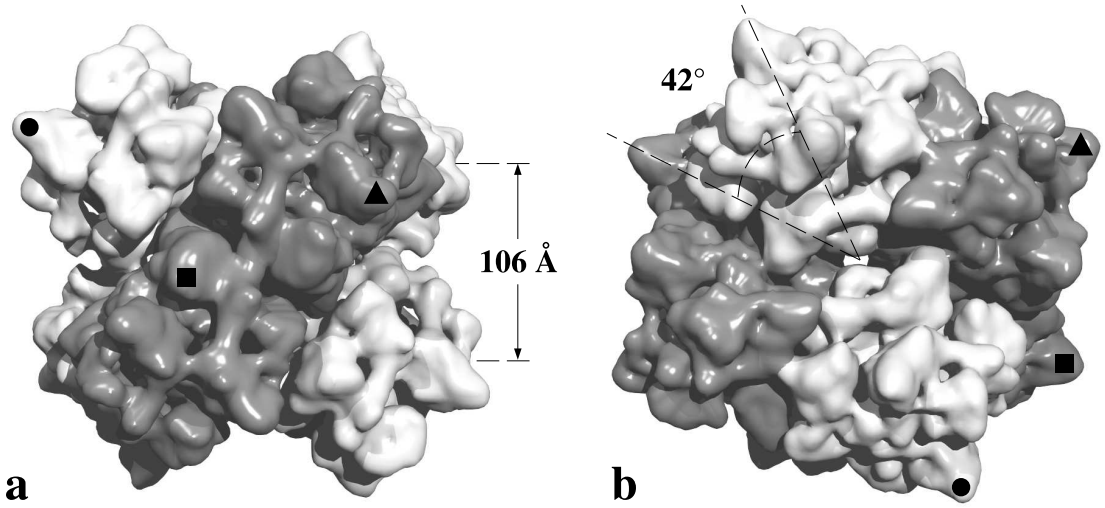


Figure 4.13: Parameters of the 8×6mer. **a** – Cross view (view along the Y axis), with the distance of the two 4×6mers indicated. **b** – Ring view (view along the Z axis), with the staggering angle between the two 4×6mers indicated. Three types of small labels mark the corners to facilitate orientation.

assembly is correct, although they had some arguments in favour of the flip-to-flip arrangement, and they clearly excluded flip-to-flop. Lamy's group collected further evidence that the flip-to-flip assembly (flop-to-flop in their terminology) is indeed the correct one ([Taveau *et al.*, 1997]). The present 10 Å cryo-EM structure confirms this assumption.

One question remained to be satisfactorily answered: why are flop-to-flop or flip-to-flop assemblies inhibited? In the electron microscope, associations of two or more 8×6mers (flop-to-flop assembly), or of additional 4×6mers to the native 8×6mer (flop-to-flop or flip-to-flop assembly) are not observed. What guarantees the integrity of the 8×6mer and blocks higher-order assemblies? Taveau *et al.* ([Taveau *et al.*, 1997]) proposed an answer for this question by overlaying stacks of the model that would correspond to a flop-to-flop and a flip-to-flop association. In both cases the same staggering angle of the native flip-to-flip association was assumed. Not allowing the stacks to rotate and assume different staggering angles leads *per se* to steric hindrances as were observed.

The 10 Å cryo-EM structure shows that the closest fit of two 4×6mers, yielding a most compact 8×6mer, results from the observed flip-to-flip assembly. This is due to the contours of V-shaped longitudinal cleft at both faces of the 4×6mer (Figure 4.12 c and Figure 4.14). At the flip face, this cleft is comparatively wide and shallow with an opening angle of 89°, and the distance between the bottom of the cleft and the edge of the 4×6mer is ∼43 Å. At the flop face, the V-shaped cleft is narrower (81°) and deeper (distance ∼51 Å). In a flip-to-flop assembly of two 4×6mers, these surface contours would not match well. Figure 4.14 shows an overlay of both clefts projected along the Y axis (cross view). In a flop-to-flop association,

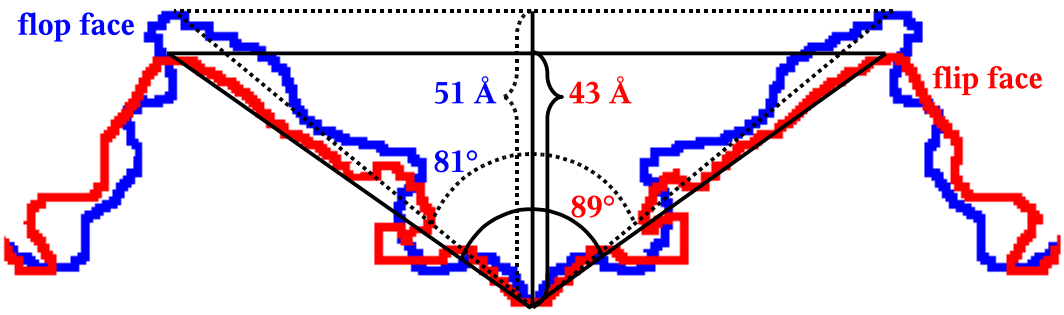


Figure 4.14: Overlay of the two V-shaped clefts of the 4×6mer projected along the Y axis (cross view). The flop side cleft is depicted in blue, the flip side cleft in red. Opening angles and depths are coloured accordingly. The overlay shows clearly the morphological differences between the two faces.

the surface contours would match well at a staggering angle of 57(±7)° and an axial distance larger than 106 Å. However, such an assembly is not observed, and would require molecular interfaces that apparently do not exist.

4.5 Molecular modelling and rigid-body fitting

Because of the high degree of sequence similarity (>75%) between chelicerate hemocyanin subunits (Figure 4.15), ternary structure modelling was straightforward on the basis of the published X-ray structure of oxygenated subunit II ([Magnus *et al.*, 1994]). For modelling of the subunits IIIA, IV, and VI, their authentic sequences were applied. These sequences have been elucidated, for the present study, by T. Burmester and his group. Because the sequences of subunits I, IIA, IIIB, and V are not yet known, structurally equivalent subunit sequences were used for modelling. In case of subunit V, the sequence of subunit VI was taken. In case of subunits I, IIA, and IIIB, the orthologous sequences of *Eurypelma* hemocyanin were applied. This is uncritical in the case of subunits I and IIA, because these subunit types (which play equivalent roles in the oligomeric architecture) are not involved in any inter-hexamer contact (Table 4.4). In the case of subunits V and IIIB, however, these circumstances should be kept in mind. Also the complete sequence of the subunit II was modelled for fitting into the density map of the cryo-EM structure. The amino acids 22–29, 134–138, 527–530, and 569–572, which are missing in the published X-ray structure of *Limulus* subunit II, were modelled by means of spatial restraints.

In the modelling procedure, 20 different homology models were calculated for each sequence. In each case, the one with the highest score in the preliminary fitting procedure with `MOLREP` into its respective isolated subunit was chosen. I considered this an objective approach of selecting a model from a number of possibilities. The

Subunit type		Contribution to interfaces		
Limulus	*Eurypelma*	2×6mer	4×6mer	8×6mer†
I°, IIA°	e*	–	–	–
II*	a*	+	–	+
IIIA*	g*	–	–	+
IIIB°	f*	–	+	+
IV*	d*	+	+	+
V°, VI*	b*, c*	+	+	–

Table 4.4: Topological correspondence of *Limulus polyphemus* and *Eurypelma californicum* subunit types. ∗ – complete amino acid sequence available; ∘ – not available; † – exclusively in *Limulus*.

chosen models were then fitted into the whole density map of the 8×6mer that was low-pass filtered to nominal 8 Å but left unmasked.

A six-dimensional grid-based exhaustive search was performed for each monomer using an angular step size of 20° and a grid spacing of 1.797 Å. The resulting monomer positions were pruned and then further refined at sub-voxel accuracy using a Powell off-lattice optimization. Finally, the whole 8×6mer was assembled from the monomer data by choosing the correct monomer for each position in the 8×6mer with the `Chimera` software (Figure 4.16). Most of the amino acids in the molecular model of the 8×6mer do not overlap. Although in some inter-hexamer interfaces a few side-chains overlap, the backbones are invariantly at proper distances. Figure 4.17 shows a comparison between projections, reprojections, surface representations and molecular model of the *Limulus polyphemus* hemocyanin.

Since the sequences were modelled by means of spatial restraints on a single subunit (X-ray structure of the oxy-form of *Limulus* subunit II ([Magnus *et al.*, 1994]), the `MODELLER` procedure ignored the presence of any neighbouring subunits. Furthermore, the orientation of the side-chains is not necessarily optimal. In case of the IV – IV bridge, a joint fitting of the two fused pdb files completely eliminated the overlaps, without moving the backbones (see below). Although further work might improve the present molecular model, it turned out to be very suitable for assessing the molecular structure of the different inter-hexamer bridges.

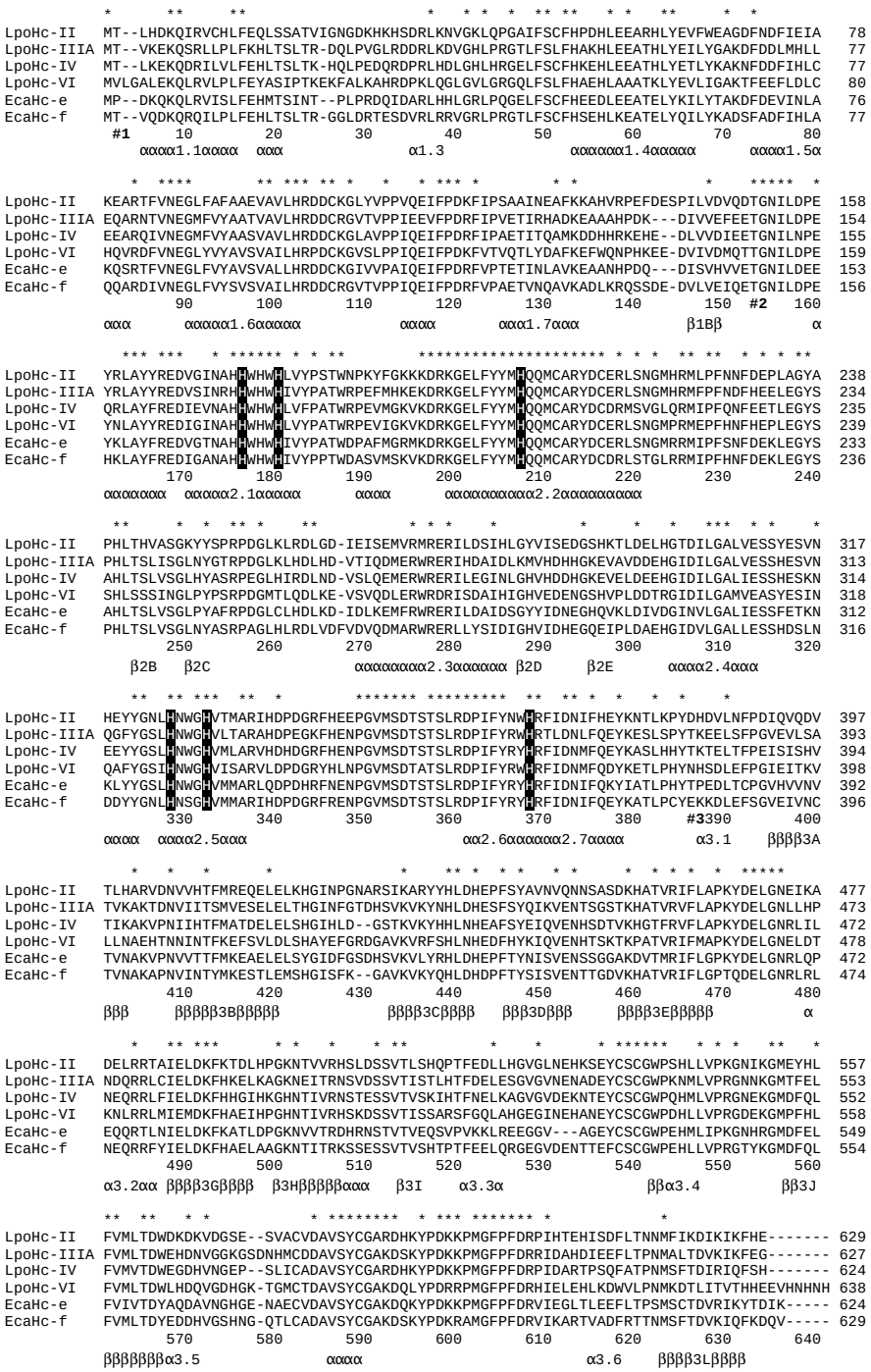


Figure 4.15: Alignment of the amino acid sequences of *Limulus polyphemus* (LpoHc) and *Eurypelma californicum* (EcaHc) hemocyanin subunits. Strictly conserved amino acids are indicated with asterisks (∗), the Cu-binding histidines are shaded in black, the secondary structure elements as deduced from the X-ray structure of subunit II ([Volbeda and Hol, 1989], [Hazes *et al.*, 1993], [Magnus *et al.*, 1994]) are indicated at the bottom. Note that the N-terminal methionine is not present in the secreted protein, as deduced from N-terminal protein sequencing [Lamy *et al.*, 1983]. #1 – domain 1, #2 – domain 2, #3 – domain 3. The alignment was kindly provided by T. Burmester.

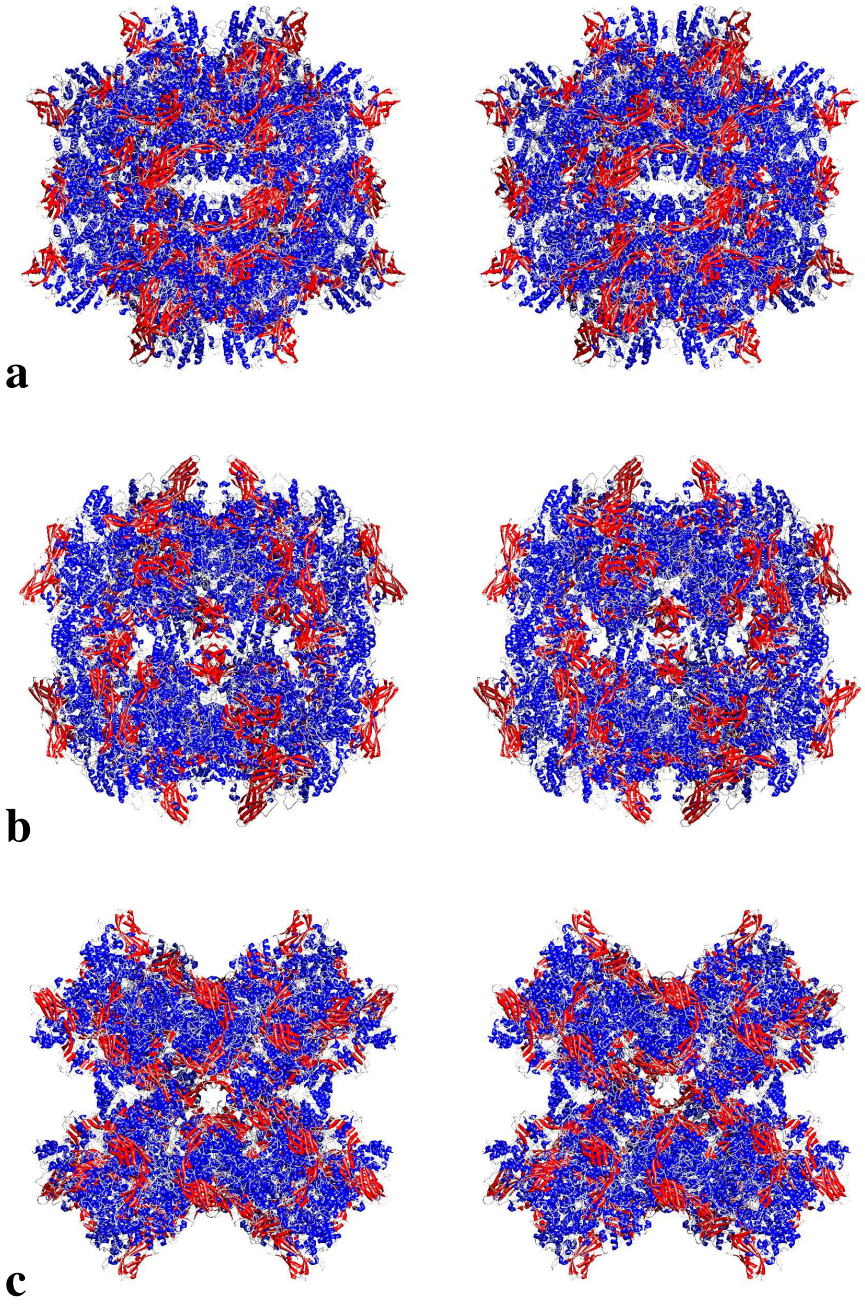


Figure 4.16: Stereo views of the molecular model of the *Limulus polyphemus* hemocyanin obtained by molecular modelling and fitting. The 48 subunits are represented as ribbons with α-helices coloured in red and β-sheets in blue. The views are oriented along the main symmetry axes and are termed as follows: **a** – Ring view (view along the Z axis), **b** – Lateral view (view along the X axis), **c** – Cross view (view along the Y axis).

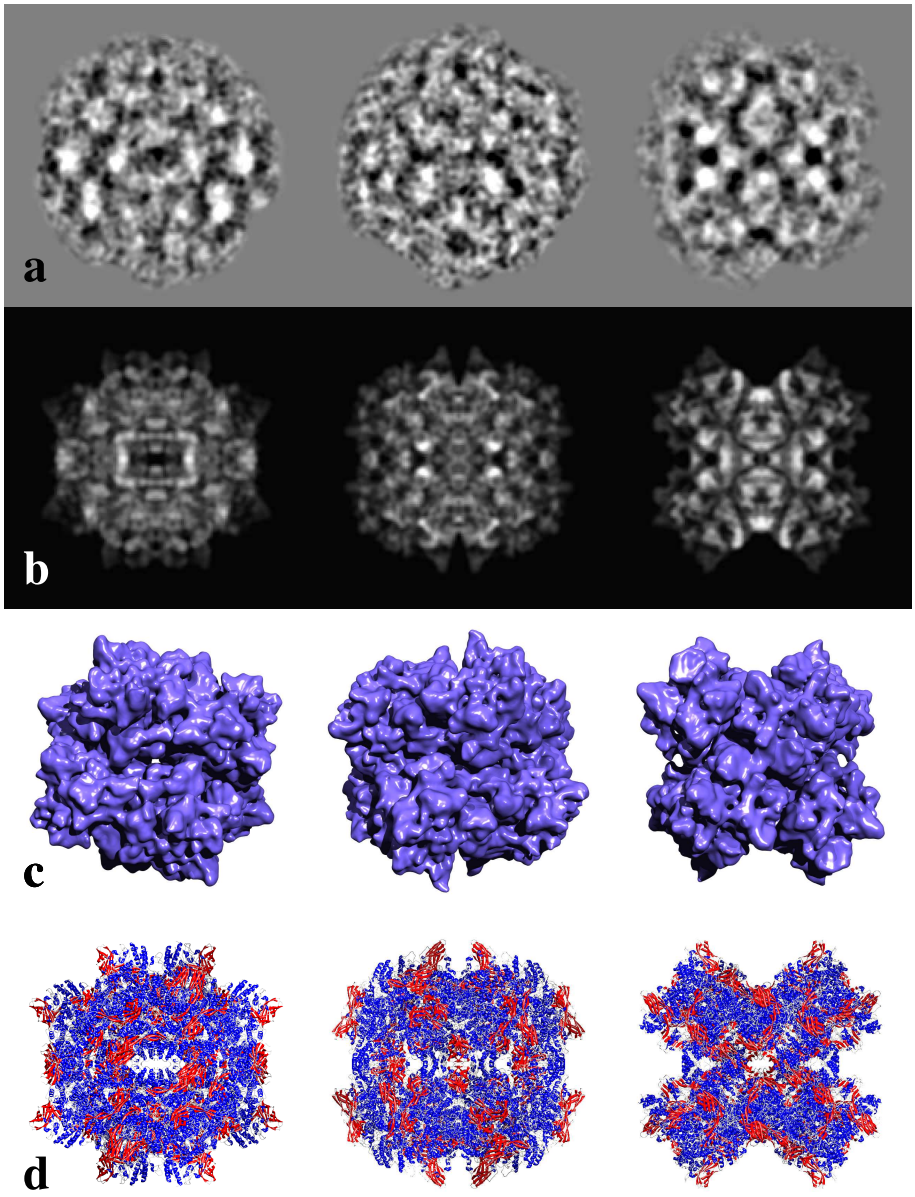


Figure 4.17: Comparison of projections, reprojections, surface representations and molecular model of the *Limulus polyphemus* hemocyanin viewed along the main symmetry axes. The columns (from the left to the right) represent views along the Z, X, and Y axis, respectively. **a** – class averages in a late stadium of the cryo-EM structure. The projections are band-pass filtered with an accent towards the high spacial frequencies to stress detail information. This procedure leads to projections that look 'noisy', but in fact the 'noise' represents highly resolved structural details; **b** – Reprojections of the final cryo-EM structure; **c** – Surface views of the final cryo-EM structure; **d** – Molecular model of the 48meric hemocyanin.

4.6 The inter-hexameric bridges

The X-ray structure of arthropod hemocyanin subunits ([Volbeda and Hol, 1989a], [Hazes *et al.*, 1993], [Magnus *et al.*, 1994]) shows three different structural domains, designated as #1, #2 and #3 (Figure 4.18). Domain #2 (red domain in Figure 4.18) contains the copper active site (turquoise spheres) that is attached to four α-helices. This domain is usually not involved in any inter-hexamer contact, in contrast to the flanking domains #1 (blue domain) and #3 (yellow domain). The N-terminal domain #1 is rich in α-helices, whereas the C-terminal domain #3 is dominated by a seven-stranded anti-parallel β-barrel and a long β-hairpin that reaches towards #1.

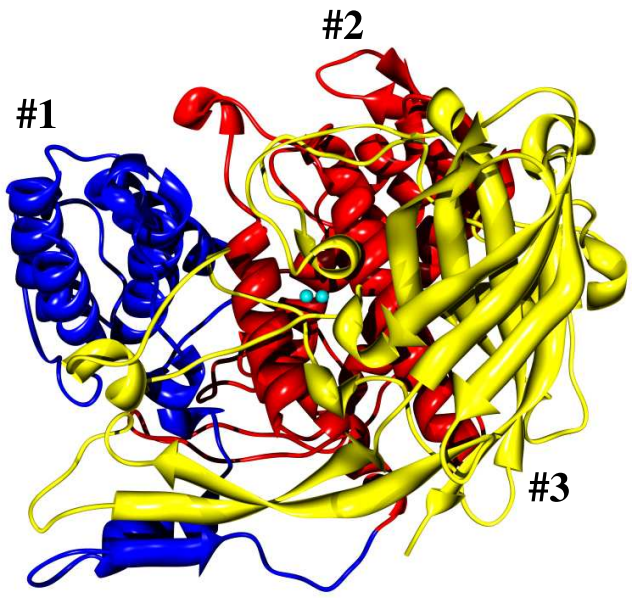


Figure 4.18: Ribbon representation of one subunit type II of the X-ray structure 1NOL of the *Limulus polyphemus* hemocyanin. The figure stresses the three different domains (domain #1 coloured in blue, domain #2 in red, and domain #3 in yellow, according to [Gaykema *et al.*, 1986]) and shows the location of the dinuclear oxygen-binding copper site (turquoise spheres). The structure represents the oxygenated T-state of the hemocyanin (molecular oxygen omitted).

As already mentioned, the molecular fitting revealed that in the 10 Å molecular model of the 8×6mer, the subunits are not exactly arranged as in the isolated spiny lobster 1×6mer, but slightly deformed, notably those that are involved in inter-hexamer contacts (see Figure 4.10). From the observation that *Limulus* 8×6mers can be dissociated into 4×6mers, 2×6mers, and subunits by changing the ionic conditions, hydrostatic bonds should be primarily expected to establish the higher-order

Contact	Topology	Amino acids of chain 1	Amino acids of chain 2
2×6mer			
II – II	outer V-edge	**Both subunits**: ^{402}HARVD^{406}(399-403), H^{410}(407), F^{412}(409), R^{414}(411), ^{436}YYHLDHEPFSY^{446}(433-443), H^{506}(503), K^{630}(625), H^{632}(627)	
II – IV	flop face	^{2}TLHD^{7}(1-4), H^{14}(11), F^{73}(70), ^{140}EFDE^{143}(137-140), H^{525}(522), H^{532}(529)	K^{404}(398), H^{410}(404), F^{412}(406), H^{441}(433), F^{444}(436), R^{476}(468), ^{612}RTPSQ^{616}(602-606), F^{631}(621), H^{633}(623)
IV – VI	flip face	K^{386}(380), E^{393}(387), E^{451}(443), H^{453}(445)	^{140}HK^{141}(139-140)
V – VI	inner V-edge	**Both subunits**: ^{26}KFALKHRDPK^{36}(25-35), ^{73}FEEFLDLCHQVRDF^{86}(72-85), Y^{92}(91)	
4×6mer			
V – V	flop face	**Both subunits**: ^{2}VL^{3}(1-2)	
V – VI	flop face	^{454}TSK^{456}(452-454), ^{568}HDQVGDH^{574}(566-572)	^{290}DEN^{292}(288-290), ^{190}EVIGKV^{195}(188-193)
VI – IIIB/IV/V	flip face	^{454}TSKTK^{458}(452-456), H^{453}(451), H^{497}(495), H^{500}(498), H^{568}(566)	**IIIB**: ^{293}EGQE^{295}(288-291), H^{287}(283), H^{291}(287)
			IV: ^{294}KE^{295}(288-289), H^{287}(281), H^{292}(286)
			V: ^{292}NGS^{294}(290-292), H^{284}(282), H^{287}(285), H^{295}(293)
8×6mer			
IIIA – IIIA	flip face	**Both subunits**: ^{140}DK^{141}(136-137), H^{138}(134), H^{430}(423)	
IIIA – IIIB	flip face	^{454}SGS^{456}(447-449), K^{574}(567)	E^{301}(297), D^{454}(450)
II – IV	flip face	D^{72}(69)	K^{71}(67)
IV – IV	flip face	**Both subunits**: F^{391}(385), ^{455}DTVKH^{459}(447-451), ^{492}FHHGIHKGH^{500}(484-492), ^{566}WE^{567}(558-559), H^{570}(562), E^{574}(566)	

Table 4.5: Table of amino acids involved in the inter-hexamer bridges. Superscript: relative residue numbers as in the main text; brackets: absolute residue numbers as in the X-ray structure of *Limulus* II and in the present molecular models of the other subunits.

quaternary structure ([Bijlholt *et al.*, 1979]). Details about the domains and amino acids involved in the inter-hexamer bridges to be discussed now are summarized in Table 4.5, the positions of the bridges are shown in Figures 4.19, 4.29, and 4.37. The α-helices and β-strands as predicted from the X-ray structure of *Limulus* subunit II are indicated in the present sequence (see Figure 4.15). The residue numbers given below conform to the numbering in this alignment, to allow the reader their clear identification. In addition to this relative numbering, the absolute residue numbers as obtained from molecular modelling are given in the Table 4.5. The code for the amino acids is listed in Table 4.6.

4.6.1 The inter-1×6mer bridges in the 2×6mer

At this structural level, the 10 Å molecular model shows four different types of interfaces that yield six bridges (Figures 4.19 and 4.20): a II – II interface with two sub-regions, a V – VI interface, two equivalent II – IV interfaces with two sub-regions, and a IV – V interface (on the flop face of the 4×6mer) with the equivalent IV – VI interface (on the flip face). Bridges between II – II, II – IV, and V – VI

Amino acid	**Abbreviation**	**Letter**
Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophane	Trp	W
Tyrosine	Tyr	Y

Table 4.6: Code for the amino acids.

have been described earlier ([Taveau *et al.*, 1997]), whereas the IV – V (respectively IV – VI) bridge is new.

The prominent II – II bridge is formed by complementary association of domain #3. In particular, both copies of the β–3A→ β–3B loop (^{402}HARVD^{406}) and the β–3C→ β–3D loop (^{440}DHEPFSY^{446}) are assembled into a rather dense structure to which other amino acids also contribute. This contrasts to the kind of association predicted in the earlier studies ([de Haas and van Bruggen, 1994], [Taveau *et al.*, 1997]). Initially, it appeared that there might be some general overlapping at this interface, but in reality the two subunits perfectly fit (apart from residues R^{404} and H^{506}/D^{406}, see below), and an intriguing contact apparatus that resembles a robot in shape (Figures 4.20 a, 4.21, and 4.22) was revealed.

The robot's shoulders are formed by two polypeptide backbones that run antiparallel and thus form a gap of 4–5 Å between eight residues (A^{403}, V^{405}, E^{442}, and P^{443}). This motif might therefore connect the two subunits via direct backbone interaction. The robot's head is an interesting association of four histidines (H^{402}, H^{506}), two serines (S^{445}), and two aspartic acids (D^{406}) arranged in a ring (this association yielded a slight overlap of H^{506} and D^{406} which requires remodelling). The head is protected by four aromatic residues (F^{444} and Y^{446}) that flank S^{445} in the sequence. Even more striking is the robot's belly. It is a chamber ∼10 Å in diameter formed by six histidines (H^{410}, H^{441}, H^{632}), two phenylalanines (F^{412}), and four acidic residues (D^{440}, E^{442}). Within this chamber, the nitrogen atoms of four basic residues (R^{404}, R^{414}) are brought together (due to this peculiar situation, the long

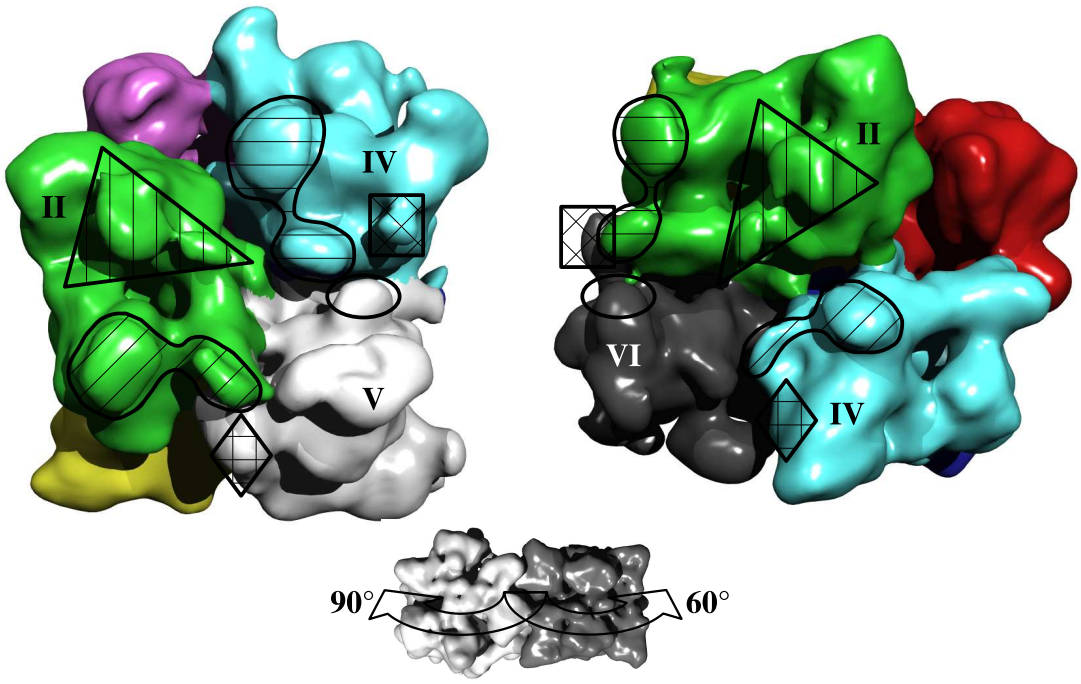


Figure 4.19: Topography of the inter-1×6mer bridges in the 2×6mer. The 2×6mer is cut through its inter-hexamer bridges, and the cut surfaces exposed to the reader, as indicated in the small grey figure at the bottom. Areas forming a joint interface (bridge) are marked by identical symbols. The interfaces connecting the 1×6mers within the basic 2×6mer are symbolized by ellipse (V – VI bridge), triangle (II – II bridge), rhombus/square (IV – V and IV – VI bridge that are structurally equivalent), and double field (II – IV bridge). Note that there are six bridges but only four types of interfaces.

residue R^{404} has not been modelled correctly and overlaps slightly with other atoms). The robot's arms are two lysines (K^{630}) that flank the chamber. The robot's feet are composed of three aromatic residues ($^{436}YYH^{438}$) that form a tripartite screen around the peptide bond of R^{414}. The two feet define the cleft between the two copies of subunit II, the opposite cleft starts above the head (see Figure 4.20 a). The chamber has two wide openings, and it is reasonable to presume that it serves the binding of an allosteric ligand. Indeed, its components are also conserved in the orthologous subunit a of *Eurypelma* hemocyanin ([Voit *et al.*, 2000]). Allosteric interaction within the 8×6mer has somehow to be transferred between the hexamers. As deduced from low resolution SAXS analyses, within the 2×6mer the rotation angle between the two hexamers is significantly increased upon deoxygenation ([Hartmann and Decker, 2002], see Table 4.1). The II – II interface right at the centre of the 2×6mer is a very promising candidate for such an allosterically controlled rotation apparatus.

The II – IV interface is localized between #1 of subunit II and #3 of subunit IV, and subdivided into two contact points. In the first one, the N-terminal end of sub-

unit II (^{2}TLHD^{7}) is connected to the β–3A→ β–3B loop of subunit IV, indicating hydrostatic bonds between a lysine (K^{404}) of subunit IV and threonine/glutamic acid of subunit II (Figures 4.20 b, 4.23, and 4.24). Interestingly, this site is surrounded by various histidines and phenylalanines and therefore might contribute to allosteric interaction between hexamers. At the second contact point, the acidic α–1.7→ β–1B loop of subunit II (^{140}EFDE^{143}) binds electrostatically to the basic-to-polar C-terminal α–helix 3.6 (^{612}RTPSQ^{616}) of subunit IV. The apolar phenylalanine F^{141} in subunit II probably shields the site from water molecules that could weaken the electrostatic bonding. Moreover, there is an opportunity for a strong salt bridge between the second glutamic acid (E^{143}) in subunit II and an arginine (R^{476}) in the β–3E→ α–3.2 loop of subunit IV (see Figure 4.20 b). Some of these secondary structures have already been predicted to be involved in this bridge ([de Haas and van Bruggen, 1994], [Taveau *et al.*, 1997]).

At the applied mass-correlated threshold there is a small inter-hexamer bridge between subunits IV and VI, and due to the D2 point-group symmetry there is an equivalent bridge between subunits IV and V. At this interface, only a few residues come together (distance ∼6 Å), namely K^{386} in α–helix 3.1 of subunit IV and ^{140}HK^{141} in the large α–1.7→ β–1B loop of subunit VI (Figures 4.20 c, 4.25, and 4.26). This association of three basic residues would require a bridging anion. However, I rather think that the positive charges prohibit bonding at this site. This is further emphasized by the fact that in the cryo-EM structure, ^{140}HK^{141} is not exactly localized within the small bridge, but protrudes sideward. I therefore think that this loop is incorrectly modelled. Tilting it towards the bridge should bring ^{140}HK^{141} close enough to two glutamic acids exposed at the opposite side of the bridge (E^{393} and E^{451}). This site would then contain two closely associated histidines (H^{140} and H^{453}), and it should be mentioned that the histidine-rich C-terminus of subunit IV is also in the neighbourhood. Whether this is of any functional significance remains to be investigated.

On the basis of the 40 Å cryo-EM structure of *Limulus* hemocyanin, it has been proposed that the V – VI bridge is constructed from five different secondary structural elements [Taveau *et al.*, 1997]. From the present 10 Å molecular model, I can reduce this to α–helix 1.5. The two copies of this helix run parallel to each other and have a backbone distance of ∼7 Å in the lower and 12 Å in the upper part (Figure 4.20 d, 4.27, and 4.28). From their orientation within this helix, the following residues potentially interact at the interface: ^{73}FEEFLDLCHQVRDF^{86}. Since the sequence of subunit V is unknown, this interface cannot be described here with certainty. With the sequence of subunit VI applied to both subunits, a few overlappings of side-chains are observed. The second element of interaction between subunit V and VI has been unpredicted, namely the sequence ^{26}KFALKAHRDPK^{36} of the large loop following helix 1.1 and leading to the short helix 1.3. It should be mentioned that helix 1.2 detected in *Panulirus* hemocyanin is missing in chelicerate hemocyanin.

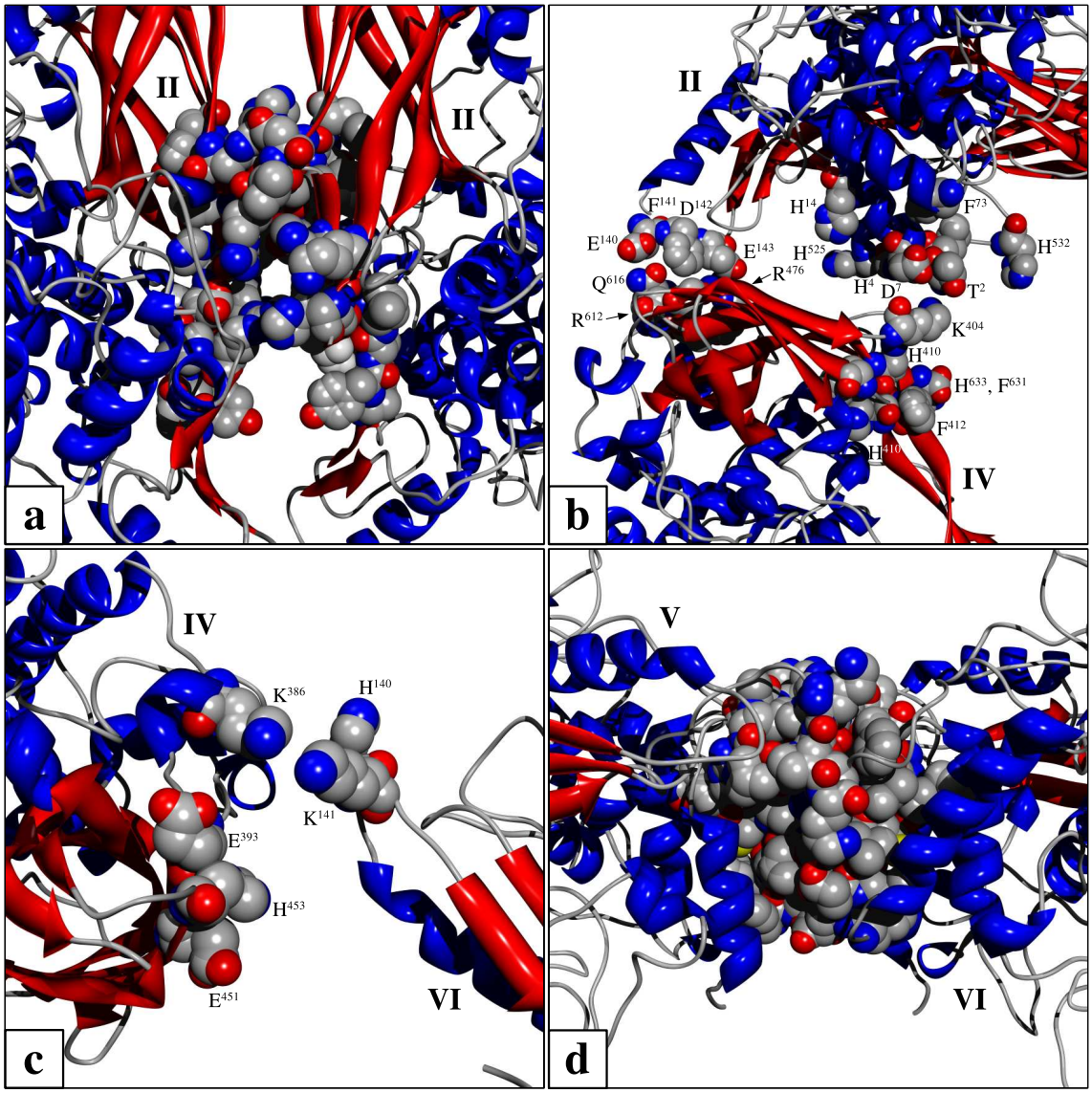


Figure 4.20: Molecular structure of the inter-1×6mer bridges in the 2×6mer. **a** – 'Robot' view of the II – II interface; **b** – The II – IV interface, note the histidine/phenylalanine clusters that are associated with the contact zone shown on the right; **c** – The IV – VI interface; **d** – The V – VI interface.

The two equivalent loops are arranged in an antiparallel manner (see Figure 4.20 d). Comparison of the loop sequence to the primary structures of the *Eurypelma* hemocyanin subunits ([Voit *et al.*, 2000]) revealed that this motif is strictly conserved in both, subunit b and c, whereas the other subunit types are highly variable in this region. In the center of this interface, a central histidine pair (H^{81}) that form the floor of a cavity with a vertical width of ~10 Å was discovered. The roof of the cavity is primarily another histidine pair (H^{32}). In the wall of the cavity, four phenylalanines (F^{27}, F^{86}) and two tyrosines (Y^{92}, in the neighboured α–helix 1.6)

are incorporated. Four additional phenylalanines (F^{73}, F^{76}) beneath the cavity surround four glutamic acid residues ($^{74}EE^{75}$). I think that like the II – II bridge, also the V – VI interface might be involved in transferring allosteric interaction between the two hexamers. Since many residues forming this interface are conserved in *Eurypelma* subunits b and c (except of H^{32}), there is a good chance that also *Limulus* subunit V possesses them, and modelling with the sequence of subunit VI might therefore be justified to describe this bridge. Nevertheless, further analysis of this interface requires the authentic sequence of subunit V.

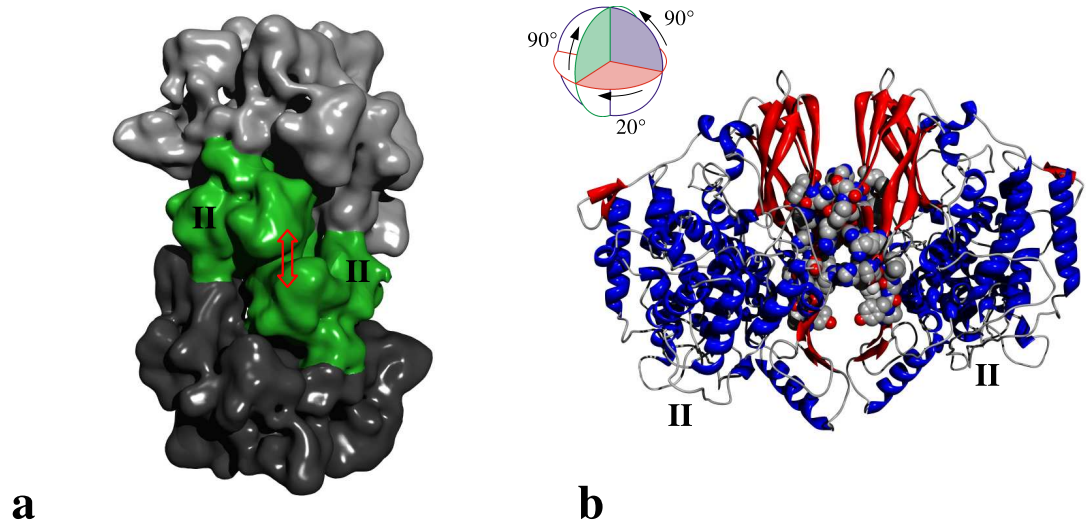


Figure 4.21: Topography and overview of the molecular structure of the II – II bridge in the 2×6mer. **a** – Surface representation of the 2×6mer with the 1×6mers coloured in light and dark grey for a better orientation. The subunits type II are highlighted in green and the double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.20 a.

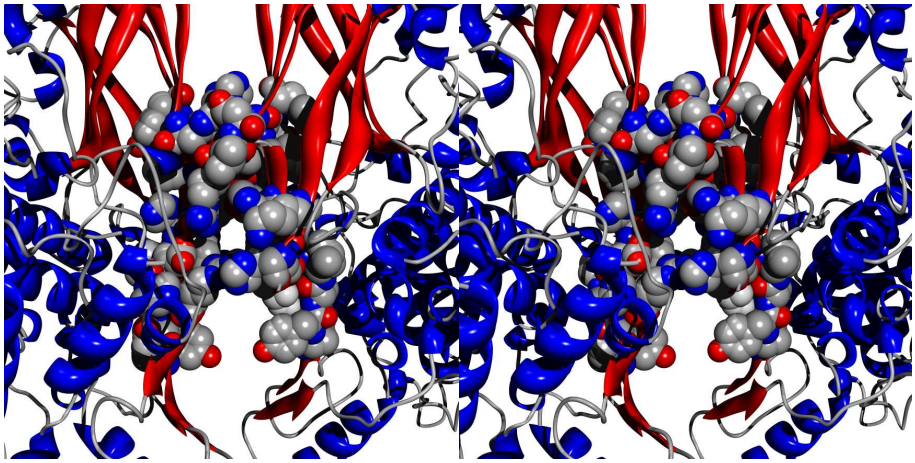


Figure 4.22: Stereo view of the molecular structure of the II – II bridge in the 2×6mer. This is the 'robot' view as described in the text. The orientation is the same as in Figure 4.20 a and Figure 4.21 b.

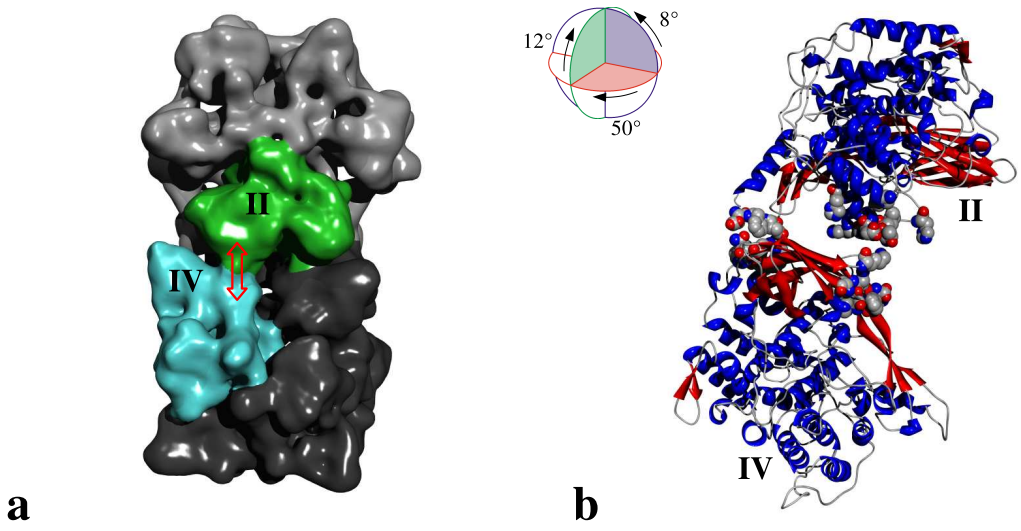


Figure 4.23: Topography and overview of the molecular structure of the II – IV bridge in the 2×6mer. **a** – Surface representation of the 2×6mer with the 1×6mers coloured in light and dark grey for a better orientation. The subunit type II is highlighted in green and the subunit type IV in turquoise. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.20 b.

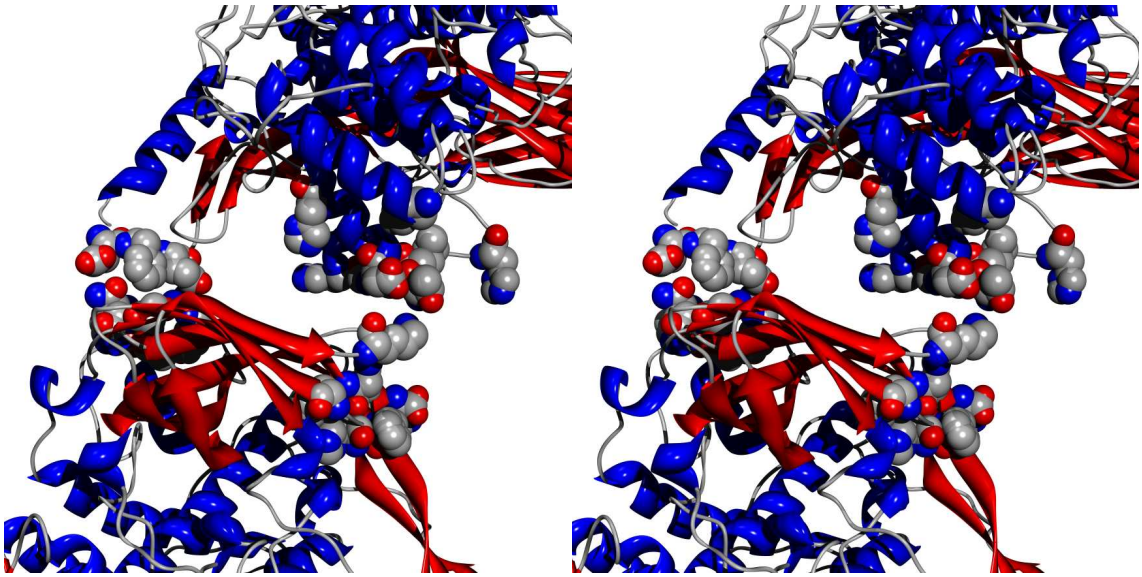


Figure 4.24: Stereo view of the molecular structure of the II – IV bridge in the 2×6mer. The orientation is the same as in Figure 4.20 b and Figure 4.23 b.

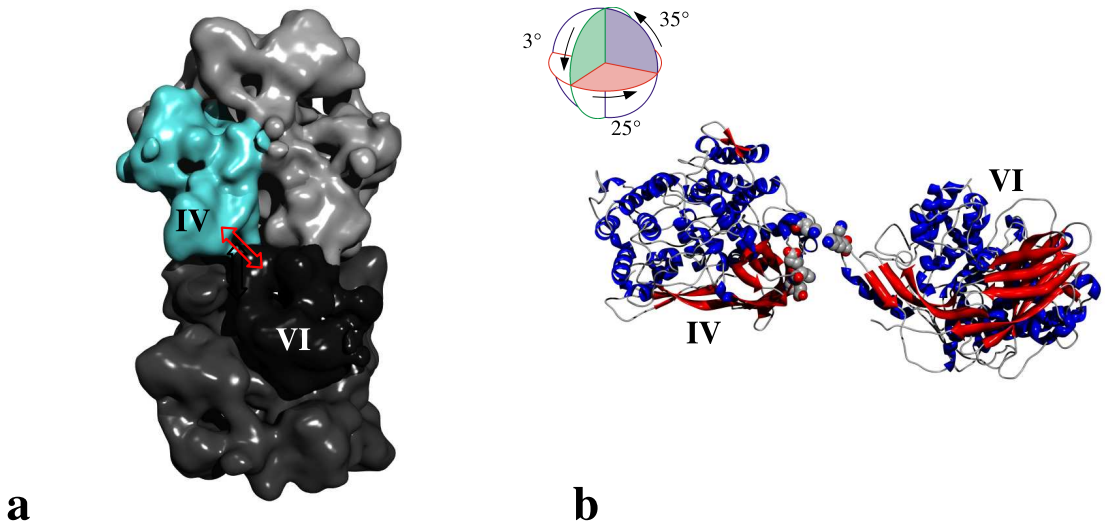


Figure 4.25: Topography and overview of the molecular structure of the IV – VI bridge in the 2×6mer. **a** – Surface representation of the 2×6mer with the 1×6mers coloured in light and dark grey for a better orientation. The subunit type IV is highlighted in turquoise and the subunit type VI in black. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.20 c.

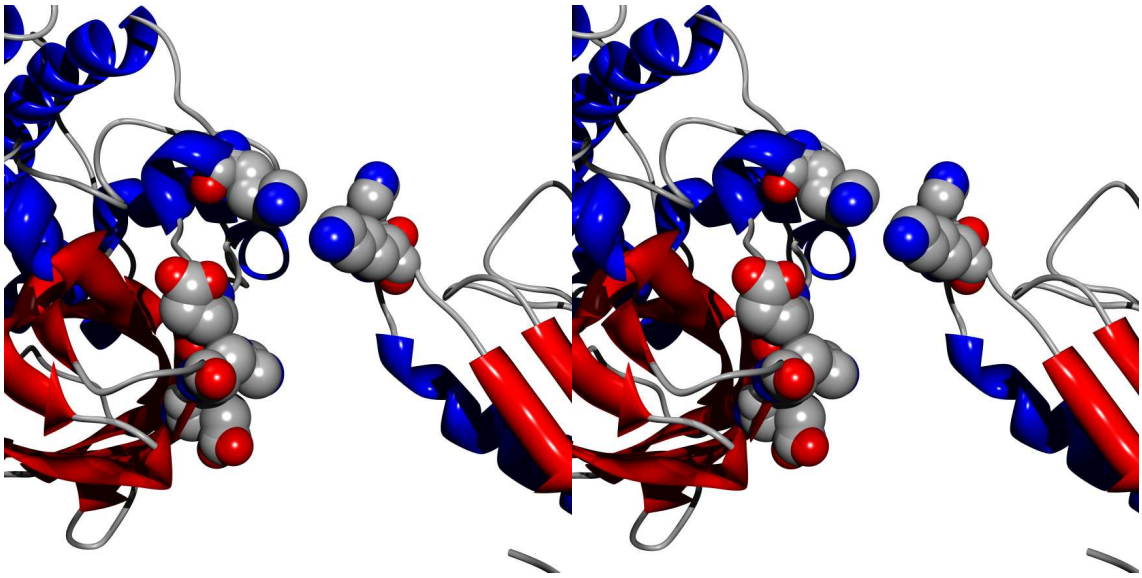


Figure 4.26: Stereo view of the molecular structure of the IV – VI bridge in the 2×6mer. The orientation is the same as in Figure 4.20 c and Figure 4.25 b.

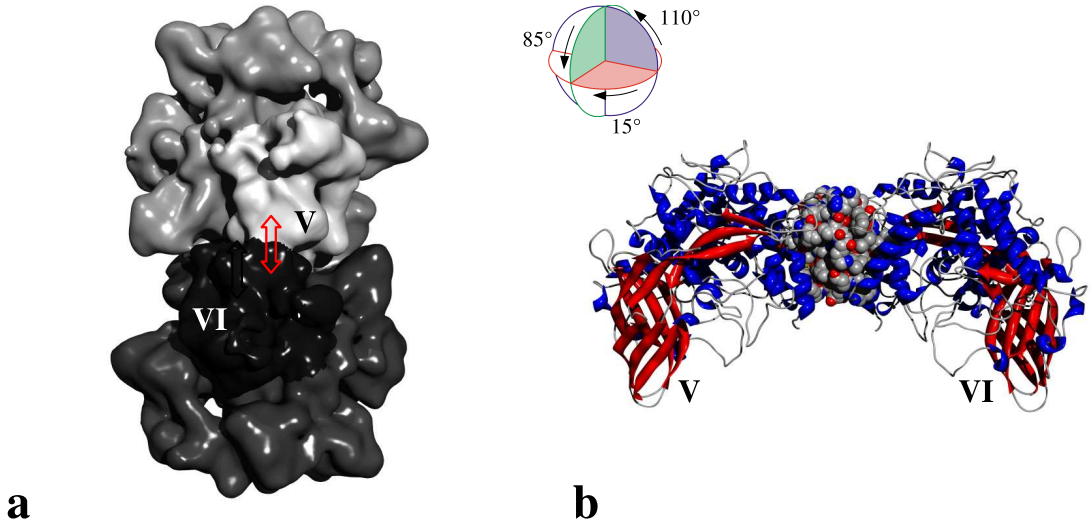


Figure 4.27: Topography and overview of the molecular structure of the V – VI bridge in the 2×6mer. **a** – Surface representation of the 2×6mer with the 1×6mers coloured in light and dark grey for a better orientation. The subunit type V is highlighted in white and the subunit type VI in black. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.20 d.

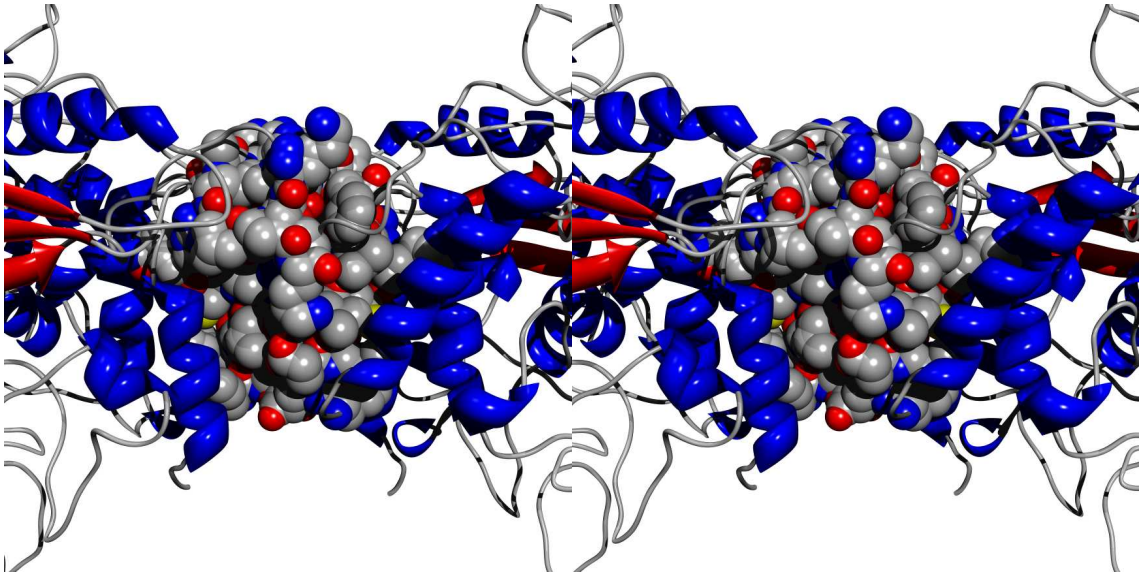


Figure 4.28: Stereo view of the molecular structure of the IV – VI bridge in the 2×6mer. The orientation is the same as in Figure 4.20 d and Figure 4.27 b.

4.6.2 The inter-2×6mer bridges in the 4×6mer

As shown in Figures 4.29 and 4.30, three types of interfaces between the two 2×6meric halves of a 4×6mer can be identified in the molecular model, yielding a total of five bridges: two equivalent VI – IIIB/IV/V interfaces at the flip face, two equivalent V – VI interfaces at the flop face, and a small V – V interface at the flop face. It should be kept in mind that the designations V and VI are convertible. It remains unknown as to which is which, and consequently, I do not know which of these two positions corresponds to the amino acid sequence of subunit VI.

On the flop face of the 4×6mer, a tiny bridge crosses the narrow cleft between the two copies of subunit V. As shown in Figures 4.30 a, 4.31, and 4.32, this bridge is due to interaction of the two N-termini (^{2}VL^{3}), these residues are identical in subunits V and VI as known from N-terminal sequencing ([Lamy *et al.*, 1983]), but the bridge cannot be formed between subunits VI because of their larger distance. The bridge is not visible, in the cryo-EM structure, at the mass-correlated threshold, but it is detectable if the threshold is slightly decreased.

The V – VI bridge is subdivided into two adjacent contact points (Figures 4.30 b, 4.33, and 4.34). In the one, the β–2D→ β–2E loop of subunit VI (^{290}DEN^{292}) joins the β–3D→ β–3E loop of subunit V (^{454}TSK^{456}) which indicates hydrostatic interaction. However, the sequence of subunit V is currently unknown and the sequence of VI has been used instead. At the other contact point, the situation is not well defined, because the large loop following α-helix 3.5 in subunit V has apparently been incorrectly modelled and protrudes from the cryo-EM structure. To form the second contact point, this loop seems to interact with its region ^{568}HDQVGDH^{574} to the region ^{190}EVIGKV^{195} of the α–2.1→ α–2.2 loop of subunit VI. An interaction between D^{573} and K^{194} is indicated, but in case of subunit V it is not the authentic primary structure. Thus, the details of this second contact remain obscure, and the whole interface requires further research. It should be noted, however, that the sequence of subunit VI provides a number of histidines in this area.

The analogous β-3D→ β-3E loop in subunit VI does not point directly towards the β-2D→ β-2E loop of subunit V, but is closer to the central hole of the opposing hexamer. There, it also joins the two other subunits of the flip face of this hexamer, namely IV and IIIB (Figures 4.30 c, 4.35, and 4.36). This VI – IIIB/IV/V interface could not be predicted from the earlier 40 Å cryo-EM structure ([Taveau *et al.*, 1997]), but is highly significant with respect to the search for centres of allosteric transmission. There is evidence from X-ray crystallography that the trimers behave as rigid bodies which can rotate 3.2° with respect to each other around their three-fold axis upon oxygenation ([Hazes *et al.*, 1993], [Magnus *et al.*, 1994]). Such a rotation would dislocate the VI – IIIB/IV/V bridge, and by domain #3 of subunit VI acting as a lever, a power transmission to the opposing 2×6mer is possible. The β-3D→ β-3E loop of subunit VI (^{454}TSKTK^{458}) joins β-strand 2E of subunit IV at ^{294}KE^{295}, and electrostatic bonding is likely to occur between them (see Fi-

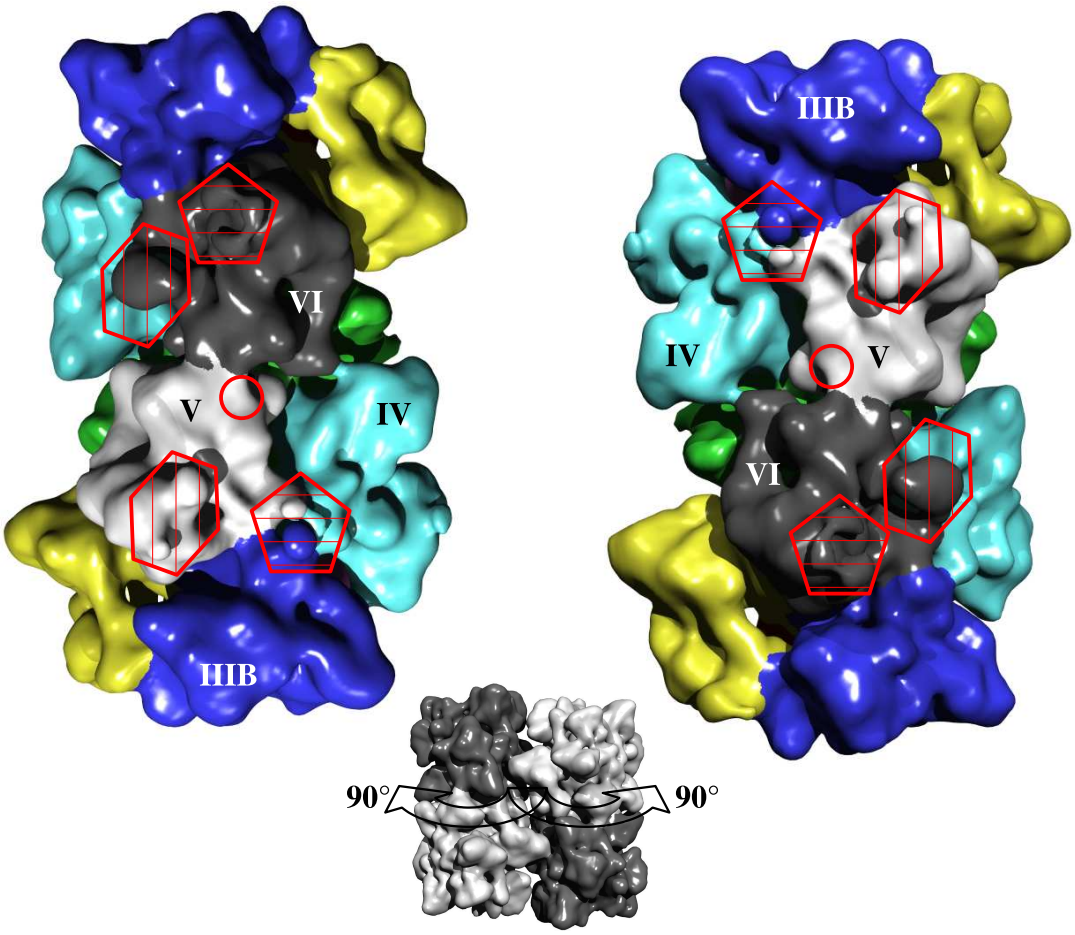


Figure 4.29: Topography of the inter-2×6mer bridges in the 4×6mer. The 4×6mer is cut through its inter-hexamer bridges, and the cut surfaces exposed to the reader, as indicated in the small grey figure at the bottom. Areas forming a joint interface (bridge) are marked by identical symbols. The interfaces connecting the 2×6mers within the 4×6mer are symbolized by cycle (V – V bridge), pentagon (VI – IIIB/IV/V bridge), and hexagon (V – VI bridge). Note that there are five bridges but only three types of interfaces.

gure 4.30 c). To this site, β-strand 2E of subunit V contributes ^{292}NGS^{294} (although this might not hold true in the authentic sequence of V), and subunit IIIB (with the orthologous sequence of *Eurypelma* subunit f applied) delivers residues ^{292}EGQE^{295} in β-strand 2E. Most interestingly, this central cluster of charged residues is surrounded by twelve histidines brought together from all four subunits (IIIB: H^{287}, H^{291}; IV: H^{287}, H^{289}, H^{292}; V: H^{284}, H^{287}, H^{295}; VI: H^{453}, H^{497}, H^{500}, H^{568}). Therefore, the VI – IIIB/IV/V interface is indeed a promising candidate for transferring allosteric interaction from one 2×6mer to the other. It differs from the V – VI bridge in two aspects: the joining of four different subunits in a single small site, and the lack of the second anchor point.

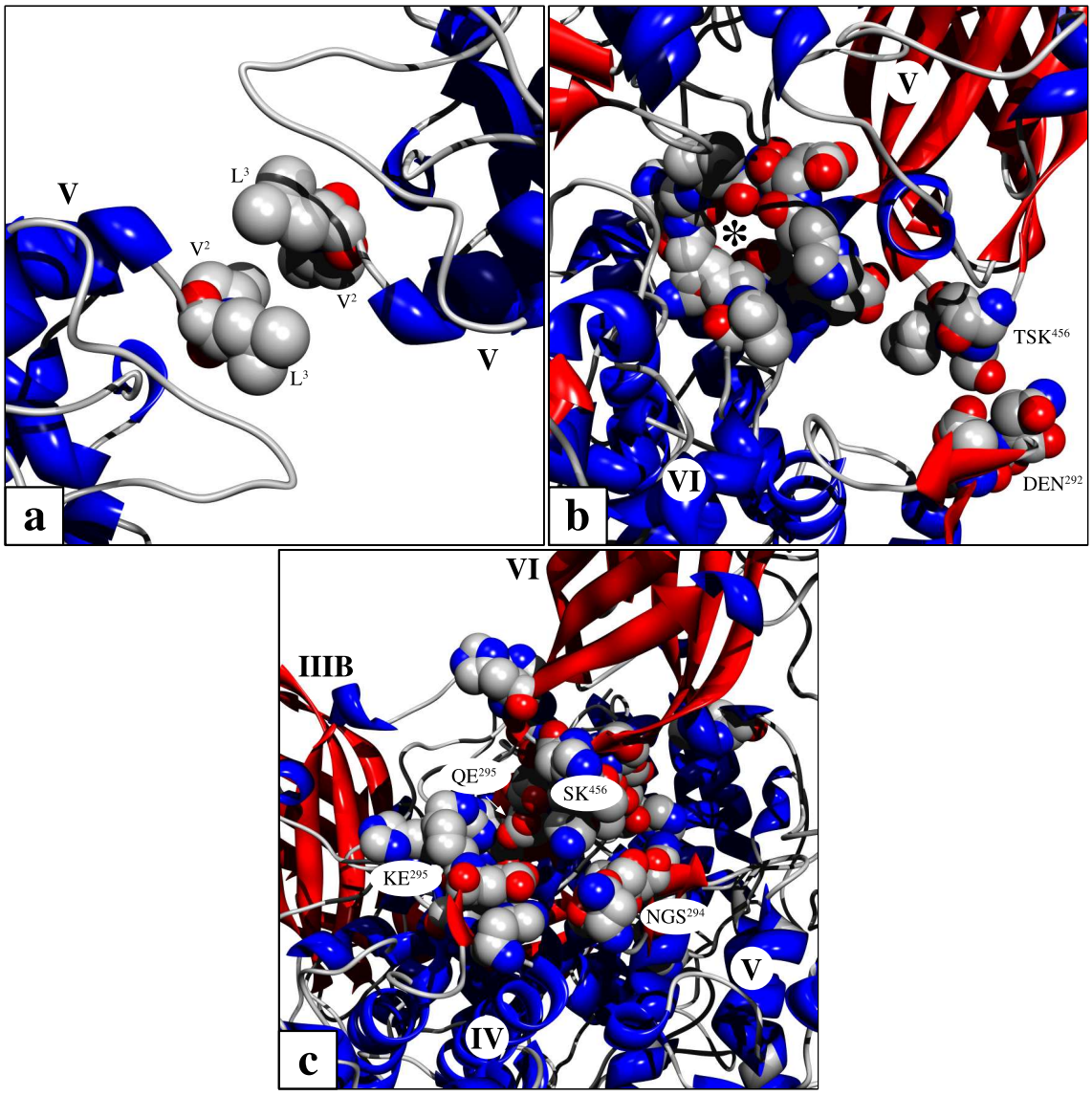


Figure 4.30: Molecular structure of the inter-2×6mer bridges in the 4×6mer. **a** – The V – V interface, formed by the N-terminal residues valine and leucine; **b** – The V – VI interface. It remains unclear which residues form the second contact (asterisk); **c** – The VI – IIIA/IV/V interface. The adjacent histidines are shown in addition to the candidate residues for chemical bonding.

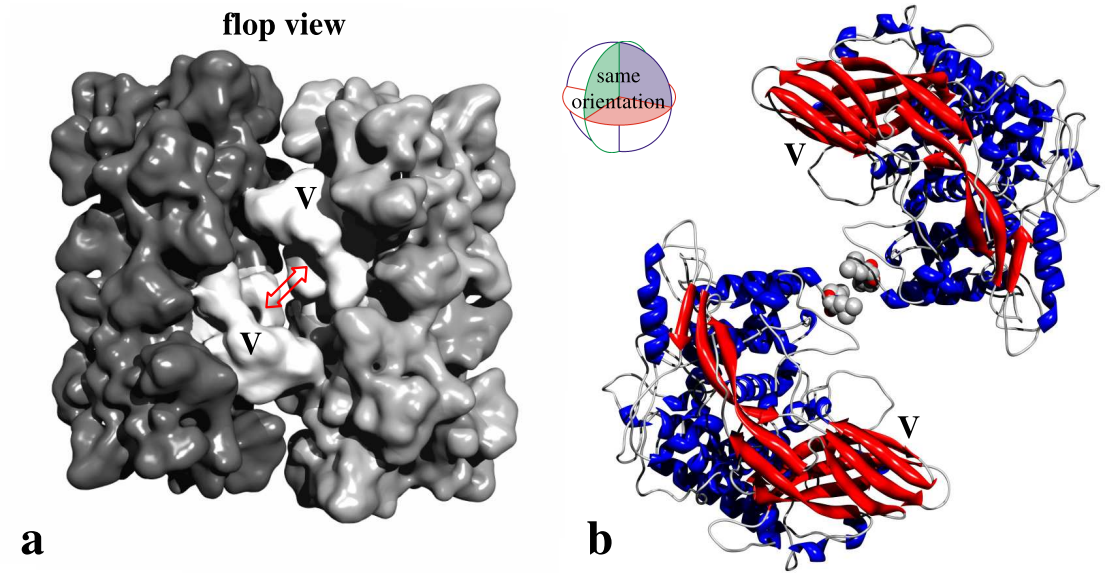


Figure 4.31: Topography and overview of the molecular structure of the V – V bridge in the 4×6mer. **a** – Surface representation of the 4×6mer with the 2×6mers coloured in light and dark grey for a better orientation. The subunit type V is highlighted in white. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) (in this case no rotation) and resembles the same orientation as in Figure 4.30 a.

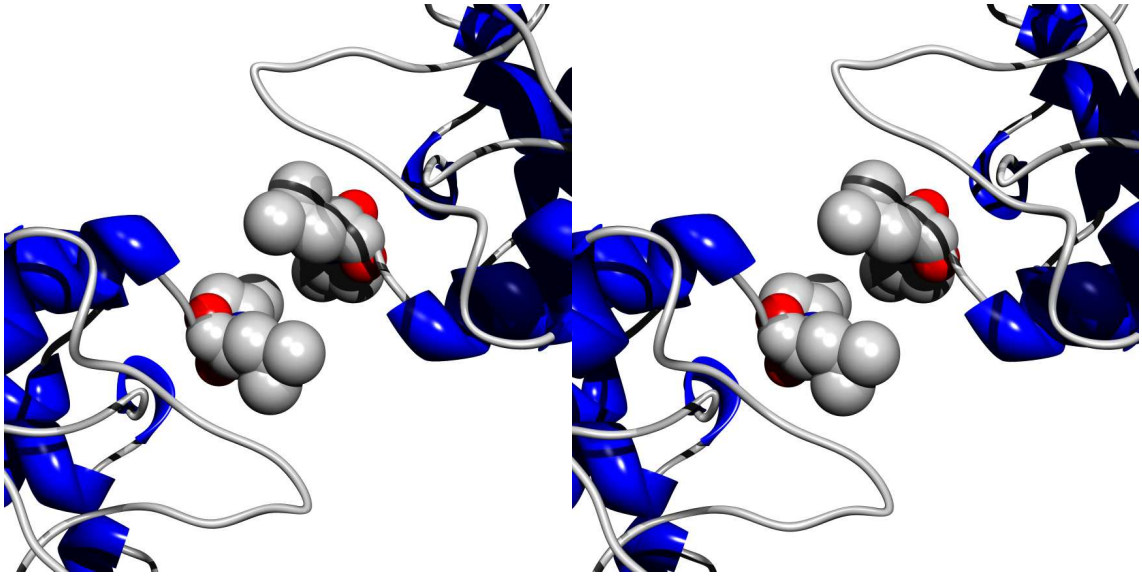


Figure 4.32: Stereo view of the molecular structure of the V – V bridge in the 4×6mer. The orientation is the same as in Figure 4.30 a and Figure 4.31 b.

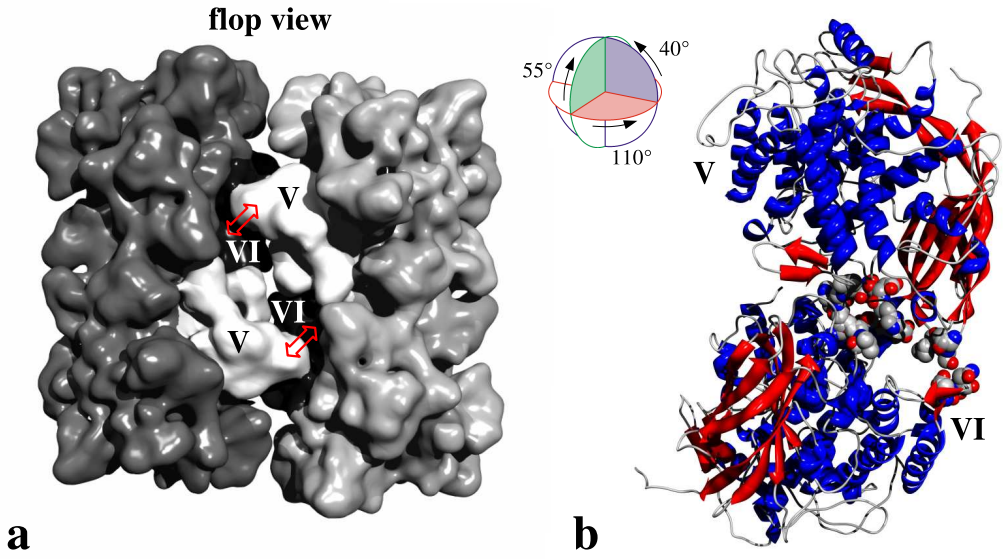


Figure 4.33: Topography and overview of the molecular structure of the V – VI bridge in the 4×6mer. **a** – Surface representation of the 4×6mer with the 2×6mers coloured in light and dark grey for a better orientation. The subunit type V is highlighted in white and the subunit type VI in black. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.30 b.

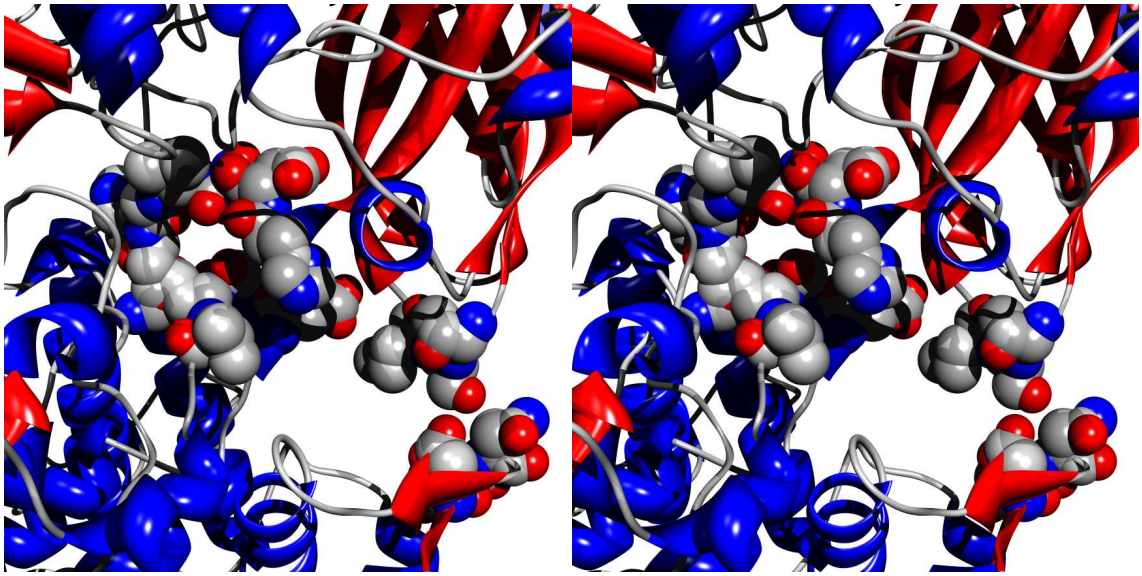


Figure 4.34: Stereo view of the molecular structure of the V – VI bridge in the 4×6mer. The orientation is the same as in Figure 4.30 b and Figure 4.33 b.

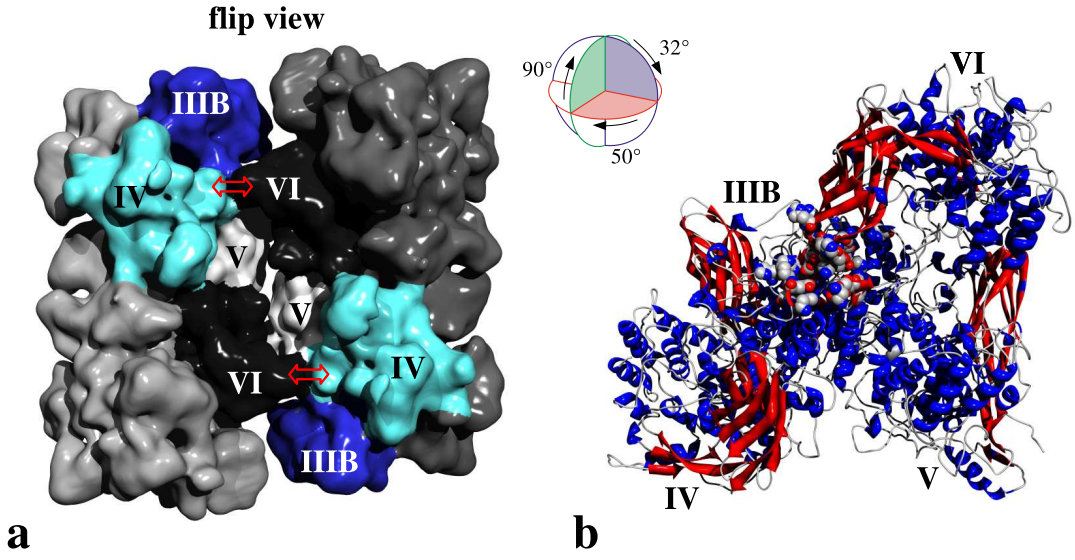


Figure 4.35: Topography and overview of the molecular structure of the VI – IIIB/IV/V bridge in the 4×6mer. **a** – Surface representation of the 4×6mer with the 2×6mers coloured in light and dark grey for a better orientation. The subunit type IIIB is highlighted in blue, the subunit type IV in turquoise, the subunit type V in white, and the subunit type VI in black. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.30 c.

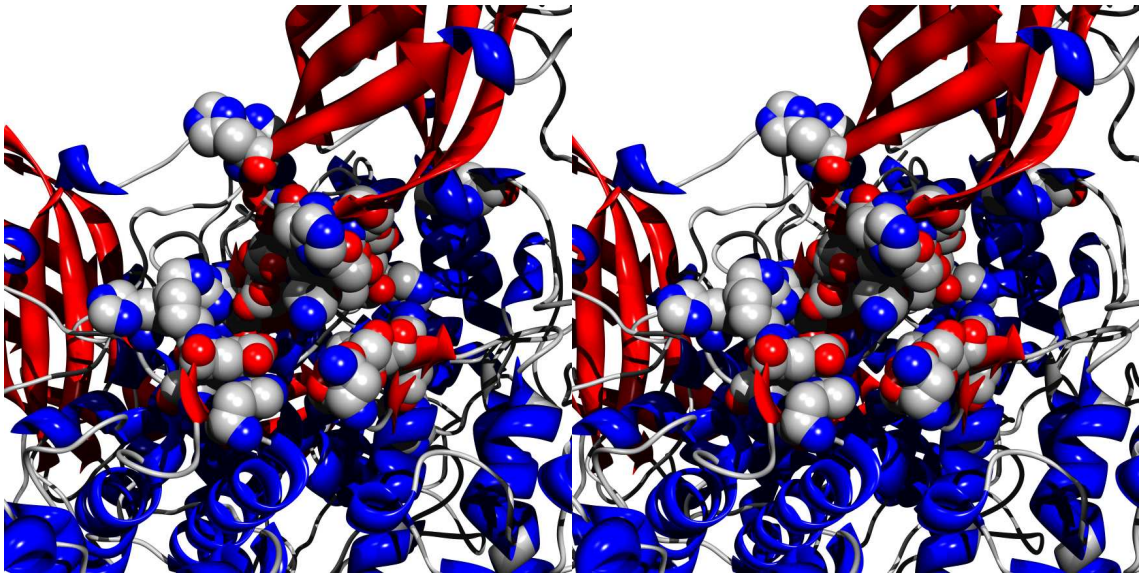


Figure 4.36: Stereo view of the molecular structure of the VI – IIIB/IV/V bridge in the 4×6mer. The orientation is the same as in Figure 4.30 b and Figure 4.35 b.

4.6.3 The inter-4×6mer bridges in the 8×6mer

Between the two 4×6mers, in the 10 Å molecular model, four different types of interfaces can be identified that together provide twelve bridges (Figures 4.37 and 4.38). Most prominent are two central, structurally equivalent IV – IV bridges. According to reassembly experiments, subunit IV is stringently required for 8×6mer formation ([van Bruggen *et al.*, 1980], [Bijlholt *et al.*, 1979], [Brenowitz *et al.*, 1983]). The other contacts are arranged in two peripheral arcs around the central IV – IV bridges (see Figure 4.37). I identified four equivalent II – IV bridges and four equivalent IIIA – IIIB bridges, and in addition two equivalent IIIA – IIIA interfaces. It is interesting to note that in reassembly experiments, in addition to subunit IV subunit IIIA also proved to be indispensable for higher-order assembly beyond the 4×6mer ([Lamy *et al.*, 1983]). Except for the IIIA – IIIA bridge, the contact regions are invariantly located on the edges of the V-shaped cleft between the two 2×6mers (see Figure 4.12 c). A IV – IV bridge as well as a IIIA – IIIB bridge have already been described ([Taveau *et al.*, 1997]), the other two interface types (II – IV and IIIA – IIIA) are new.

The IIIA – IIIA bridge is easily detected in the cross view representation of the cryo-EM structure and is formed between #1 and #3 (Figures 4.38 a, 4.39, and 4.40). The fitted ternary structures revealed a most interesting interface. It is formed by two copies of ^{140}DK^{141} in the α–1.7→ β–1B loop that suggest two strong salt bridges between the two subunits. The distance between the two potential binding atoms of the fitted molecular model is 4.5 Å. This arrangement is closely associated with four histidines (H^{138} in the same loop and H^{430} in the β–3B→ β–3C loop), their aromatic rings occupy the corners of a tetrahedron with edge length 8 Å. Viewed from the top, the four histidines expose a wide opening, probably for a ligand (see Figure 4.38 a). Due to its peripheral position and oblique orientation between the two 4×6mers, chemical chances within the two IIIA – IIIA bridges (that are localized at opposite surfaces of the 8×6mer) might result in a slight rotation of the two 4×6mers with respect to each other around their staggering axis. This could transmit forces over long distances and thereby mediate allosteric interaction between the two 4×6mers. The intricate molecular structure of the IIIA – IIIA interface is very promising with respect to this role.

The small IIIA – IIIB bridges are not visible in the cryo-EM structure at the mass-correlated threshold used, but appear if the threshold is slightly decreased. In the case of IIIB, the orthologous sequence of *Eurypelma* hemocyanin (subunit f) was applied. However, since *Eurypelma* hemocyanin is a native 4×6mer, it could be expected that regions involved in inter-4×6mer bridges differ in the two sequences. Also earlier authors ([Taveau *et al.*, 1997]) described a IIIA – IIIB bridge, but their predicted sites differ completely from those found in the present study. I detected potential linkages between both #3, and between #3 of subunit IIIA and #2 of subunit IIIB (Figures 4.38 b, 4.41, and 4.42). To the #3 – #3 contact, IIIA contributes lysine K^{574} in the loop following helix α–3.5. To the #3 – #2 contact it delivers

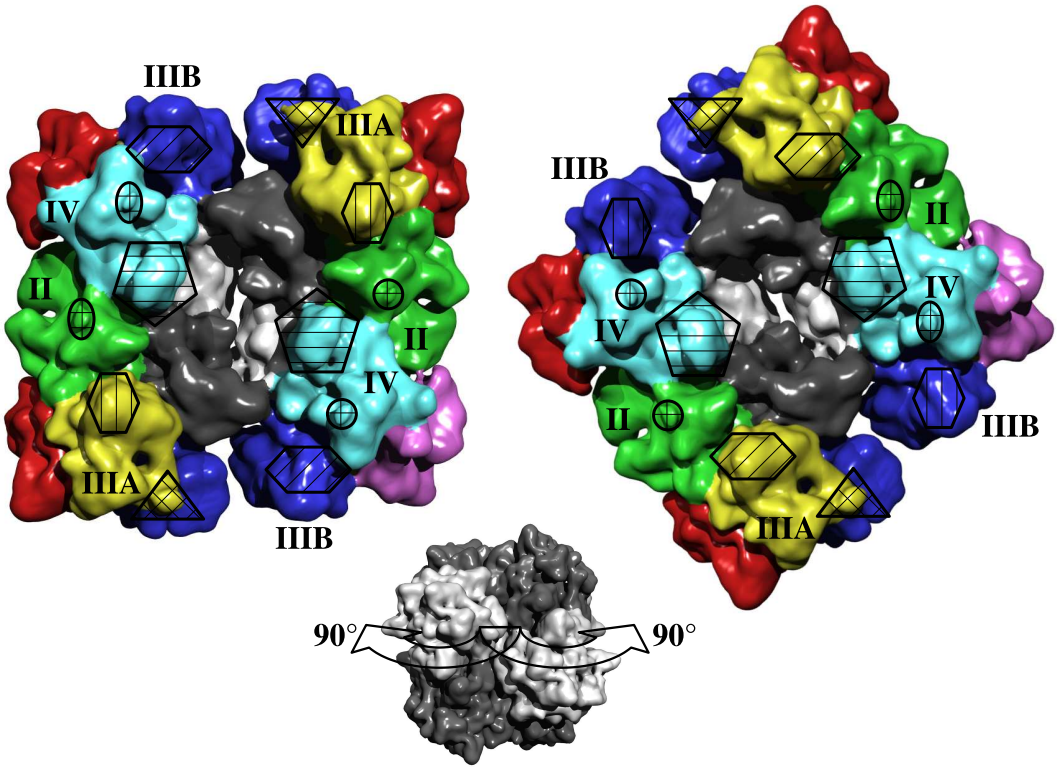


Figure 4.37: Topography of the inter-4×6mer bridges in the final 8×6mer. The 8×6mer is cut through its inter-hexamer bridges, and the cut surfaces exposed to the reader, as indicated in the small grey figure at the bottom. Areas forming a joint interface (bridge) are marked by identical symbols. The interfaces connecting the 4×6mers within the 8×6mer are symbolized by cycle/ellipse (II – IV bridge), triangle (IIIA – IIIA bridge), pentagon (IV – IV bridge), and hexagon (IIIA – IIIB bridge). Comparison to Figure 4.12 a, a' visualizes that the 4×6mers are assembled at their flip faces. Note that there are twelve bridges but only four types of interfaces.

the sequence ^{454}SGS^{456} in the β–3D→ β–3E loop. This indicates hydrostatic bonds in both cases. Indeed, with the orthologous *Eurypelma* type f sequence applied for IIIB, at a proper distance acidic residues (D^{454} in the β-3D→ β-3E loop, respectively E^{301} in the β–2E α–2.4 loop) are available for bonding. However, the correct description of this bridge requires the authentic primary structure of subunit IIIB.

The II – IV bridge is localized, in both subunits, at exactly the same position in the α–1.4→ α–1.5 loop (Figures 4.38 c, 4.43, and 4.44). There is ionic bonding possible between residue D^{72} of subunit II and residue K^{71} of subunit IV, all other possibilities can be excluded from the distances. As seen in Figure 4.38 c, the lysine residue points into the wrong direction, but this might be due to incorrect modelling. The backbone distance between the two amino acids is ∼10 Å which is easily bridged if the two residues would be oriented towards each other.

In the prominent, somewhat oblique IV – IV bridge, three regions of #3 are involved (see Table 4.5), which in principle have already been traced in the earlier 40 Å cryo-EM structure ([Taveau *et al.*, 1997]). Now, they could be more firmly

identified as ^{455}DTVKH^{459} in the β–3D→ β–3E loop, ^{492}FHHGIHKGH^{500} from β–strand 3G to β–strand 3H, probably phenylalanine F^{391}, and several amino acids in α–helix 3.5 and the following loop (^{566}WE^{567}, H^{570}, and E^{574}). Although this loop is missing in the X-ray structure of *Limulus* II hemocyanin, I found very few overlaps of side-chains in this region. These overlaps could be eliminated by a joint remodelling of the two subunits (see above) which did not influence the backbone, but re-orientated several side-chains and resulted in a perfect fit at the interface.

Despite the numerous amino acids involved in this site, a clear – and fascinating – picture emerges (Figures 4.38 d, 4.45, and 4.46). This interface contains the following elements: the relevant secondary structures from both subunits are arranged in an antiparallel manner, and their equivalent turns ^{455}DTVKH^{459} join each other at 4 Å distance. This assembles the two hydrophobic valines (V^{457}) that apparently help to bind the two turns together. This binding might be reinforced by hydrostatic interaction between the charged residues D^{455}, T^{456} and K^{458} delivered from both subunits. More peripheral at either flank of the interface, bonding is likely to occur between residues K^{498} and E^{567}, as well as between H^{500} and E^{574}. The clue of this interface, however, is two equivalent residue clusters which resemble minute 'wheelhouses': they are located at either side of the joined turns. The 'wheel' is the central C6-ring of a tryptophan (W^{566}), and the 'house' is four closely associated histidines (H^{459}, H^{493}, H^{497}, H^{570}) that surround the tryptophan. An additional histidine (H^{494}) and two phenylalanines (F^{391} and F^{492}) are in the neighbourhood (at 8–10 Å distance from the tryptophan), but it is not clear whether they contribute to this site. More significant is probably that one of the four central histidines (H^{497}), with a ring-to-ring distance of only 2 – 3 Å to the tryptophan, comes from the opposite subunit. If the tryptophan would be dislocated with respect to H^{497} by an allosteric effect, this could transmit forces between the opposite backbones to change their distance. This, in turn, would change the length of the IV – IV bridge, and due to its oblique orientation, this would change the staggering angle between the two 4×6mers. It therefore appears that the two IV – IV bridges (together with the two IIIA – IIIA bridges between the alternating hexamers, see Figure 4.37) have the potential to transmit allosteric interaction between the 4×6mers.

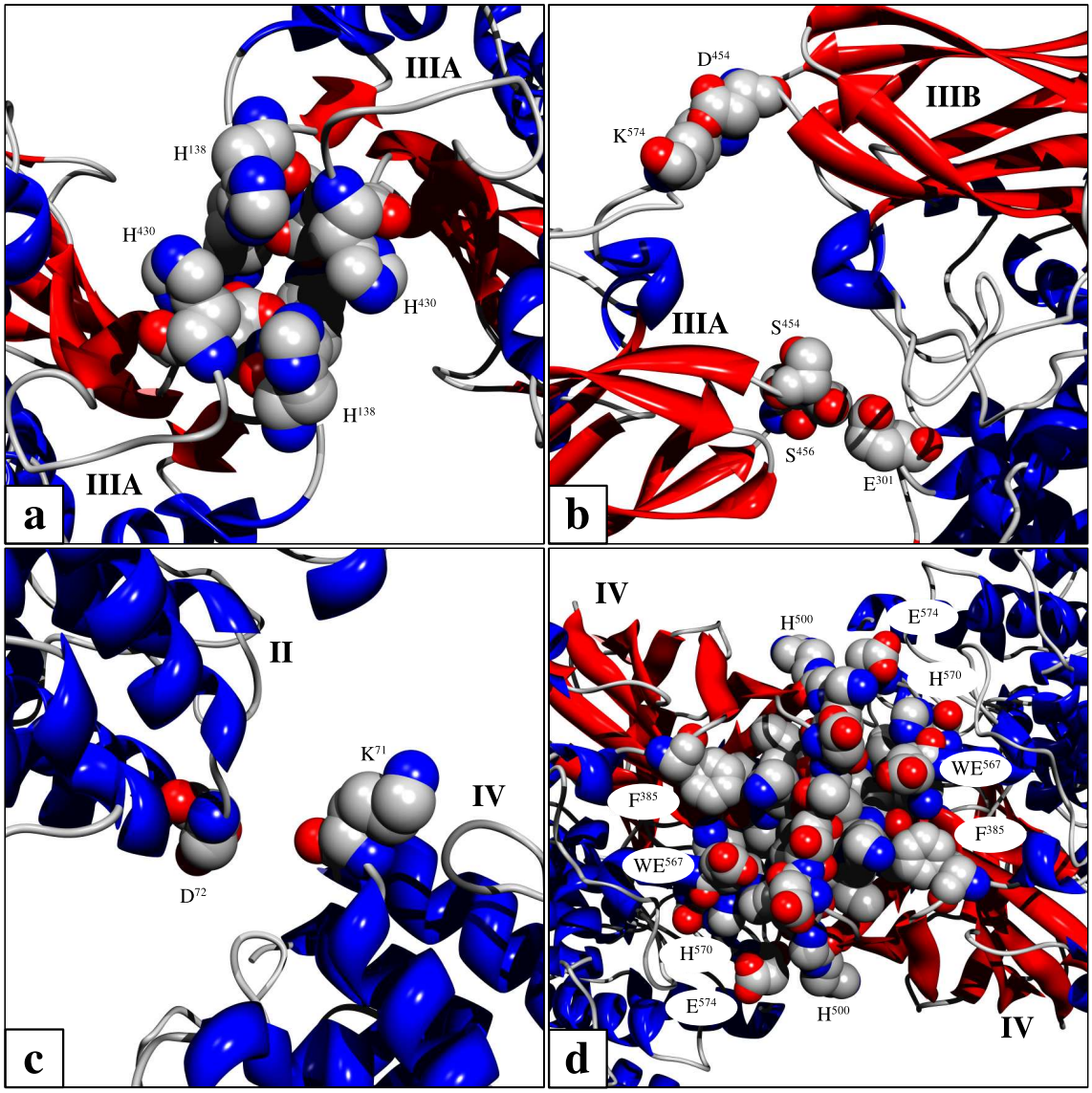


Figure 4.38: Molecular structure of the inter-4×6mer bridges in the 8×6mer. **a** – The IIIA – IIIA interface; **b** – The IIIA – IIIB interface; **c** – The II – IV interface; **d** – The IV – IV interface. This is the 'wheelhouses' view as described in the text.

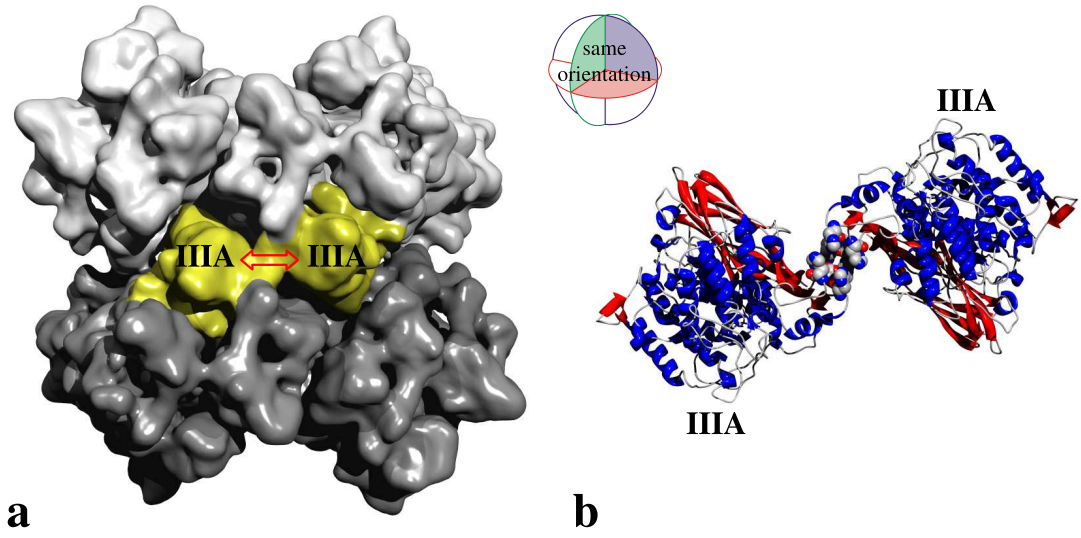


Figure 4.39: Topography and overview of the molecular structure of the IIIA – IIIA bridge in the 8×6mer. **a** – Surface representation of the 8×6mer with the 4×6mers coloured in light and dark grey for a better orientation. The subunits type IIIA are highlighted in yellow. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) (in this case no rotation) and resembles the same orientation as in Figure 4.38 a.

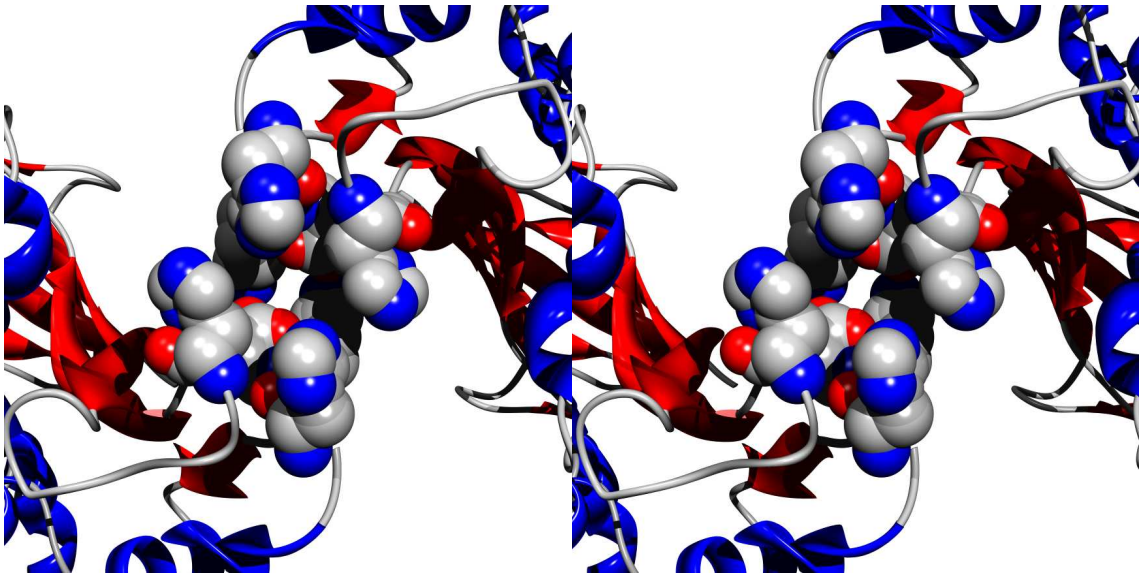


Figure 4.40: Stereo view of the molecular structure of the IIIA – IIIA bridge in the 8×6mer. The orientation is the same as in Figure 4.38 a and Figure 4.39 b.

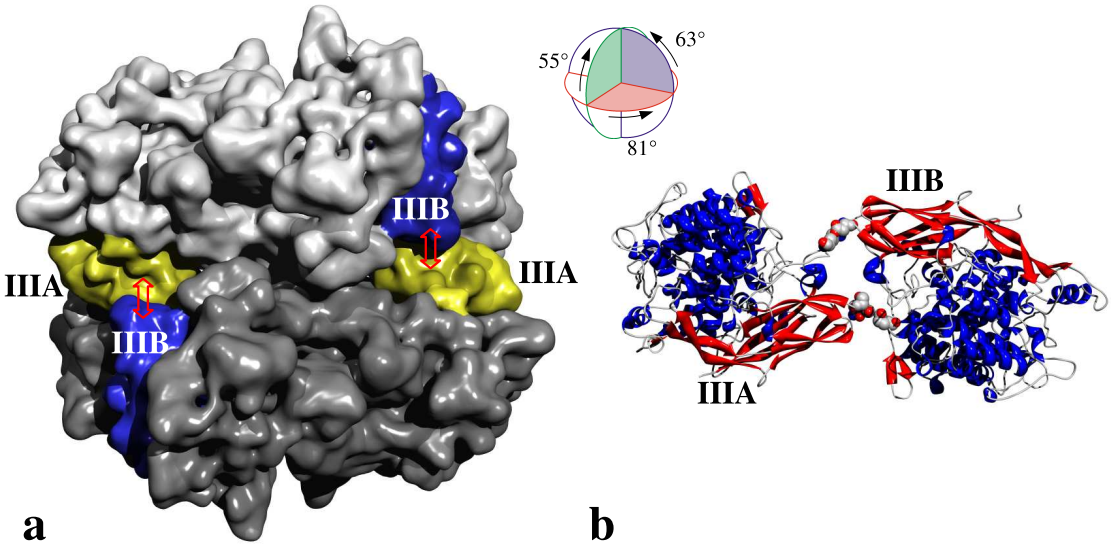


Figure 4.41: Topography and overview of the molecular structure of the IIIA – IIIB bridge in the 8×6mer. **a** – Surface representation of the 8×6mer with the 4×6mers coloured in light and dark grey for a better orientation. The subunits type IIIA are highlighted in yellow, the subunits type IIIB in blue. The double arrows indicate the positions of the bridges; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.38 b.

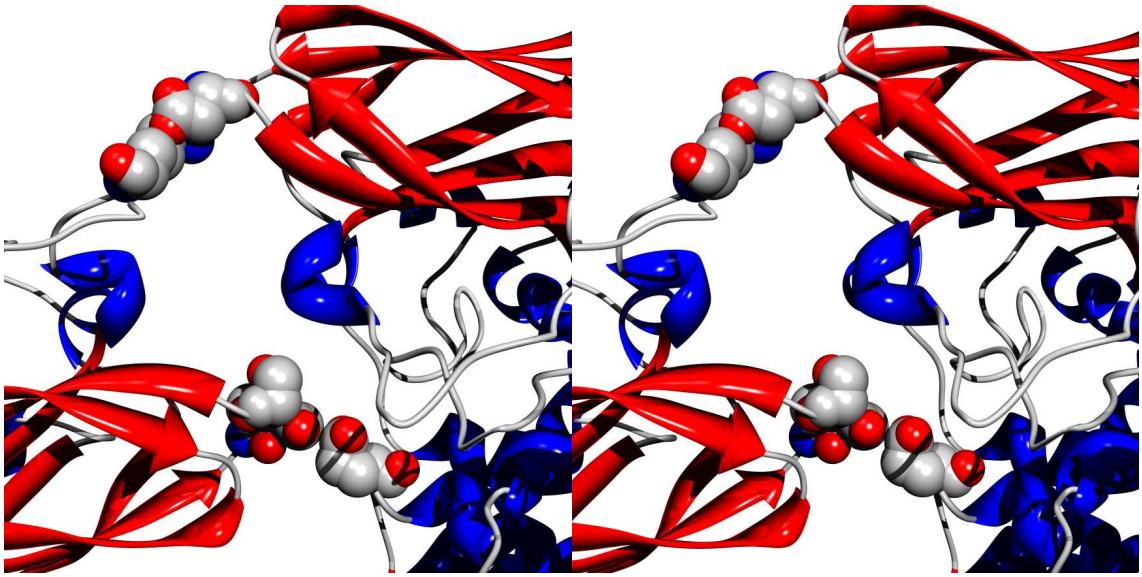


Figure 4.42: Stereo view of the molecular structure of the IIIA – IIIB bridge in the 8×6mer. The orientation is the same as in Figure 4.38 b and Figure 4.41 b.

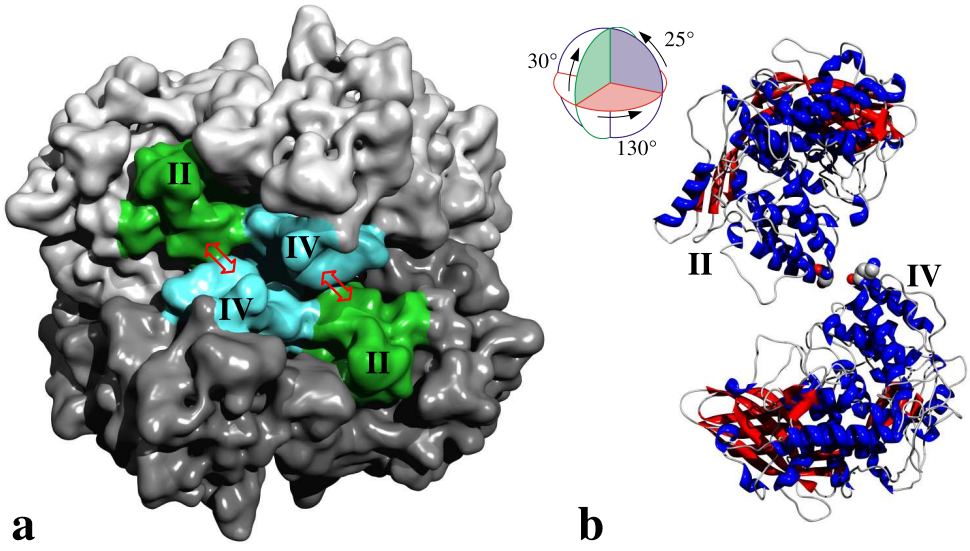


Figure 4.43: Topography and overview of the molecular structure of the II – IV bridge in the 8×6mer. **a** – Surface representation of the 8×6mer with the 4×6mers coloured in light and dark grey for a better orientation. The subunits type II are highlighted in green, the subunits type IV in turquoise. The double arrows indicate the positions of the bridges; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.38 c.

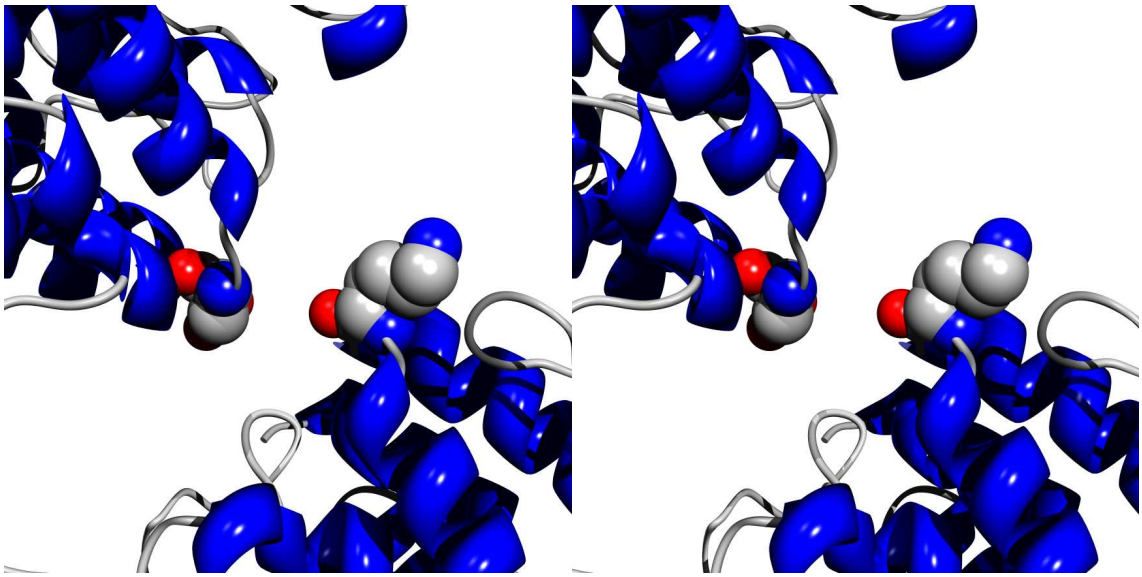


Figure 4.44: Stereo view of the molecular structure of the II – IV bridge in the 8×6mer. The orientation is the same as in Figure 4.38 c and Figure 4.43 b.

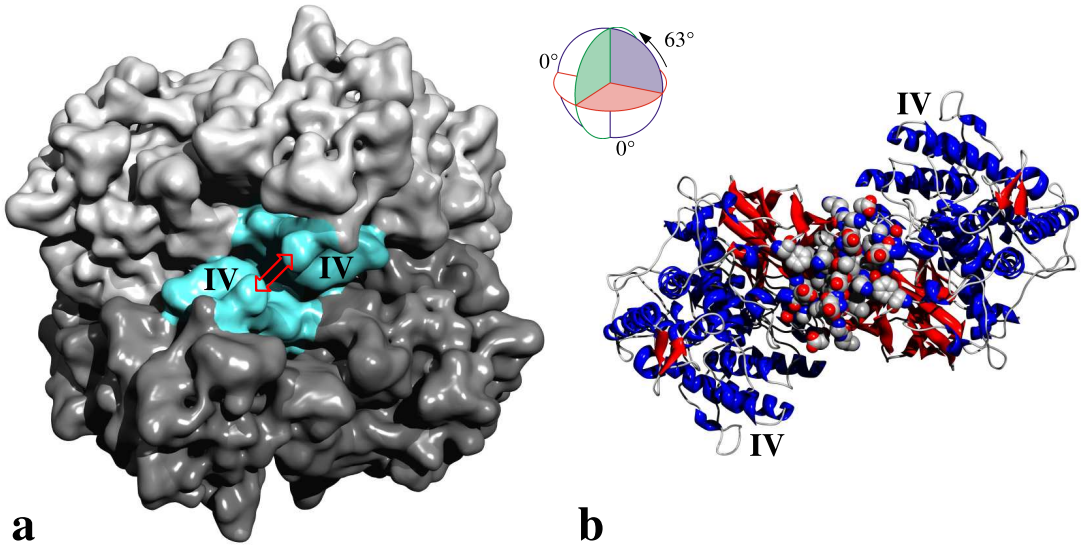


Figure 4.45: Topography and overview of the molecular structure of the IV – IV bridge in the 8×6mer. **a** – Surface representation of the 8×6mer with the 4×6mers coloured in light and dark grey for a better orientation. The subunits type IV are highlighted in turquoise. The double arrow indicates the position of the bridge; **b** – Overview of the molecular structure of the bridge. The sphere on the top indicates the rotation of the molecular model in relationship to the surface representation in (a) and resembles the same orientation as in Figure 4.38 d.

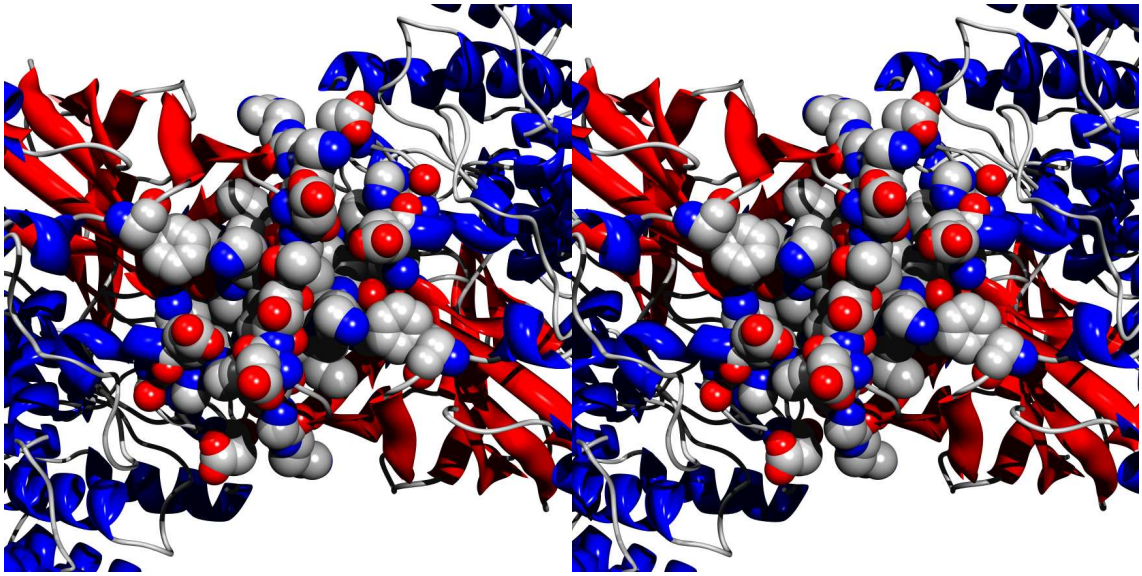


Figure 4.46: Stereo view of the molecular structure of the IV – IV bridge in the 8×6mer. The orientation is the same as in Figure 4.38 d and Figure 4.45 b.

4.7 Stereochemical quality of the models

The molecular models chosen to create the entire 8×6mer have just few residues with 'disallowed' bondig angles and the amount is comparable to amounts found in X-ray structures (e.g. 1OXY). The total numbers of residues in 'disallowed' regions are three for model EcaHc_e_01 (*Limulus* subunit types I and IIA, Figure 4.47), two for model LpoHc2_20 (*Limulus* subunit type II, Figure 4.48), six for model LpoHc3A_20 (*Limulus* subunit type IIIA, Figure 4.49) and model EcaHc_f_19 (*Limulus* subunit type IIIB, Figure 4.50), respectively, seven for model LpoHc4_19 (*Limulus* subunit type IV, Figure 4.51), and nine for model LpoHc6_06 (*Limulus* subunit types V and VI, Figure 4.52). The residues are listed in Table 4.7. Some of the residues in 'disallowed' regions are present in the contact regions. They are designated with asterisks in Table 4.7 and summarized in Table 4.8 with their respective contacts.

Subunit type	Model name	Residues in 'disallowed' regions
I/IIA	EcaHc_e_01	$R^{30}(25)$, $E^{152}(144)$, $N^{573}(561)$
II	LpoHc2_20	$I^{25}(23)$, $D^{152}(150)$
IIIA	LpoHc3A_20	$R^{24}(21)$, $Q^{27}(23)$, $D^{34}(30)$, $K^{133}(129)$, $H^{138}(134)^*$, $E^{152}(145)$
IIIB	EcaHc_f_19	$S^{33}(29)$, $V^{35}(31)$, $E^{152}(147)$, $S^{193}(188)$, $F^{267}(262)$, $T^{534}(527)$
IV	LpoHc4_19	$T^{23}(20)$, $K^{24}(21)$, $E^{30}(26)$, $D^{31}(27)$, $K^{133}(129)$, $E^{152}(146)$, $K^{404}(398)^*$
V/VI	LpoHc6_06	$K^{24}(23)$, $L^{29}(28)^*$, $D^{34}(33)^*$, $K^{36}(35)^*$, $K^{133}(132)$, $E^{142}(141)$, $T^{152}(150)$, $I^{156}(154)$, $T^{407}(404)$

Table 4.7: Table with the residues of the models in 'disallowed' regions of the Ramachandran plot according to the software PROCHECK. Amino acids involved in the contact regions are designated with asterisks (∗). Superscript: relative residue numbers as in the main text; brackets: absolute residue numbers as in the X-ray structure of *Limulus* II and in the present models of the other subunits.

Molecular dynamics calculations are required to improve the quality of the molecular models. This calculations were not performed for two reasons. First, the quality of the molecular models created by means of spatial restraints is comparable to the quality of X-ray structures. Second, most of the affected residues are not involved in the contact regions. Even though there are five 'problematic' residues in contacting areas (see Table 4.8), the relaxation of their bond angles will not move them out of their respective contact regions. Consequently, the interpretation of the contact regions presented above are not affected by the residues in 'disallowed' regions. Second, the molecular models were created upon the X-ray structure 1NOL ([Magnus *et al.*, 1994]). It is assumed that this structure represents the oxygenated

Subunit type	Model name	Residues	Contact
IIIA	LpoHc3A_20	$H^{138}(134)$	8×6mer IIIA – IIIA
IV	LpoHc4_19	$K^{404}(398)$	2×6mer II – IV
V/VI	LpoHc6_06	$L^{29}(28)$, $D^{34}(33)$, $K^{36}(35)$	2×6mer V – VI

Table 4.8: Table with the 'disallowed' residues in contact regions and their respective contacts. Superscript: relative residue numbers as in the main text; brackets: absolute residue numbers as in the X-ray structure of *Limulus* II and in the present models of the other subunits.

T-state of the hemocyanin from *Limulus polyphemus*. Prior the very time expensive molecular dynamics calculations, molecular models upon the X-ray structure 1HCY ([Volbeda and Hol, 1989a]) of the hemocyanin from *Panulirus interruptus* should be calculated. This structure represents the deoxygenated R-state of the protein. Only when correlation coefficients elucidated which is the most probable state of the cryo-EM structure, either T- or R-state, then refinements of the molecular models should be calculated.

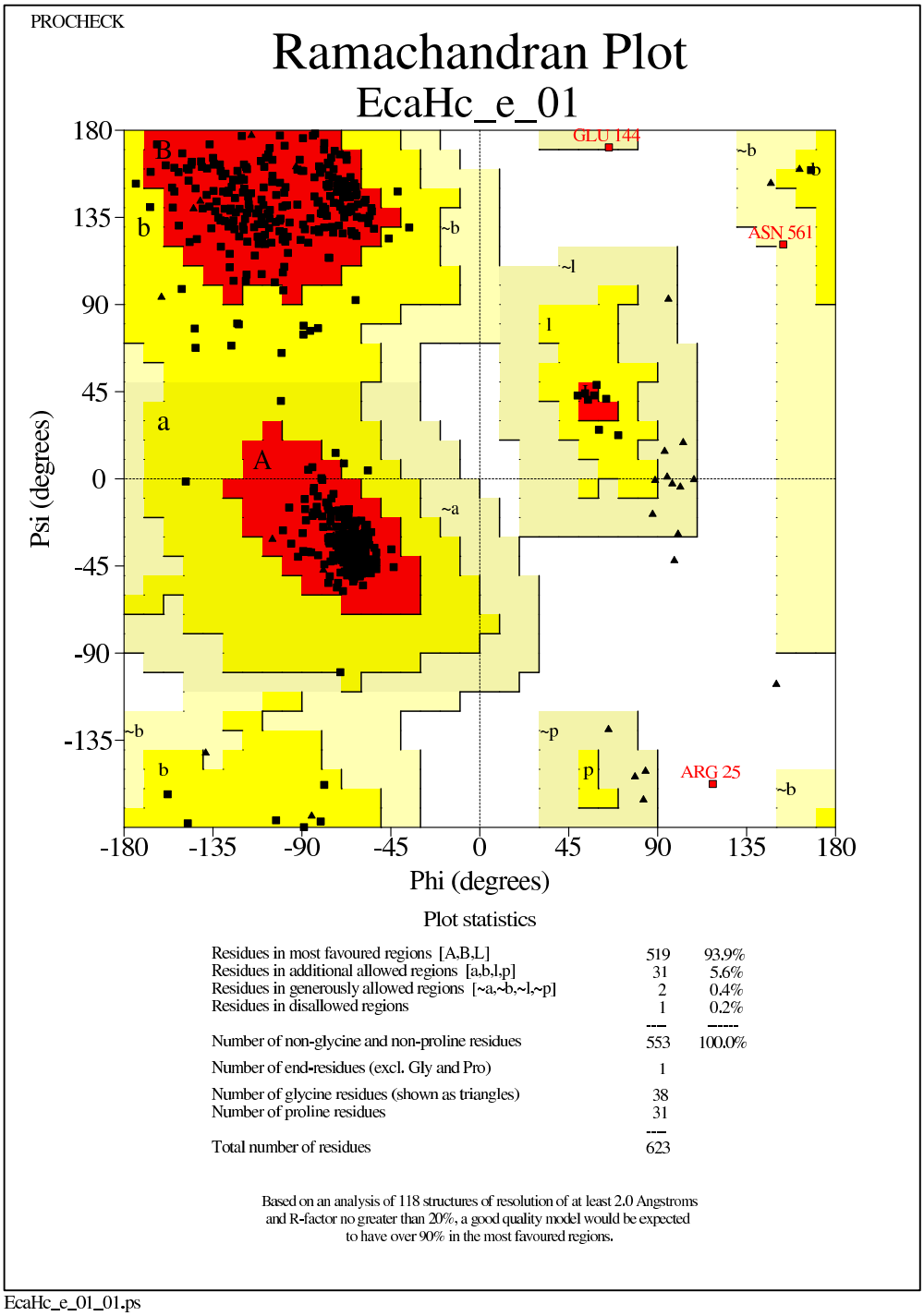


Figure 4.47: Ramachandran plot of the molecular model EcaHc_e_01 representing *Limulus* subunit types I and IIA. Residues in 'disallowed' regions are $R^{30}(25)$, $E^{152}(144)$, $N^{573}(561)$. White areas – disallowed regions; cream areas – generously allowed regions (∼a, ∼b, ∼l, ∼p); yellow areas – additional allowed regions (a, b, l, p); red areas – most favoured regions (A, B, L); triangles – glycine residues; black squares – non-glycine residues; red squares – residues in 'disallowed' regions

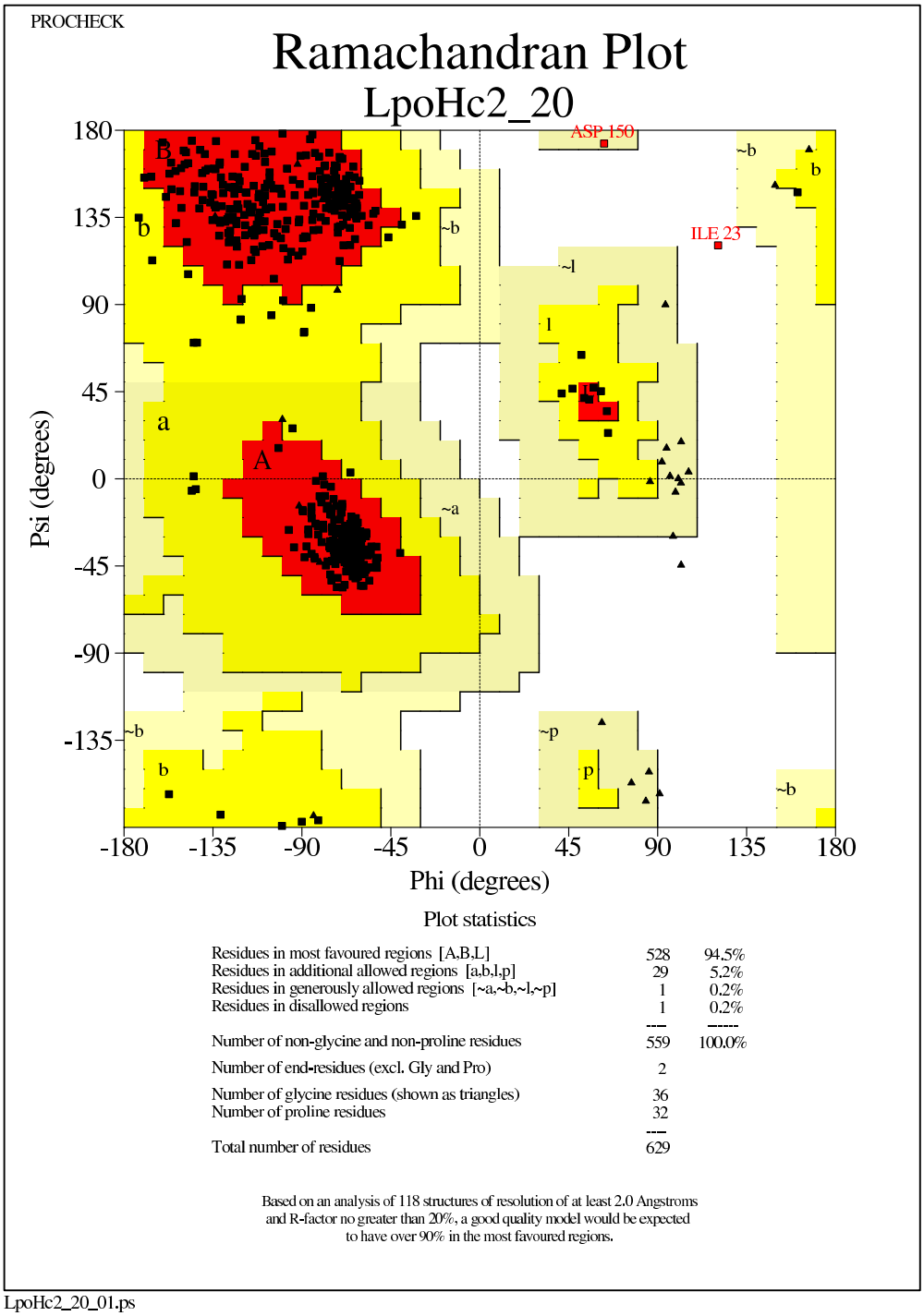
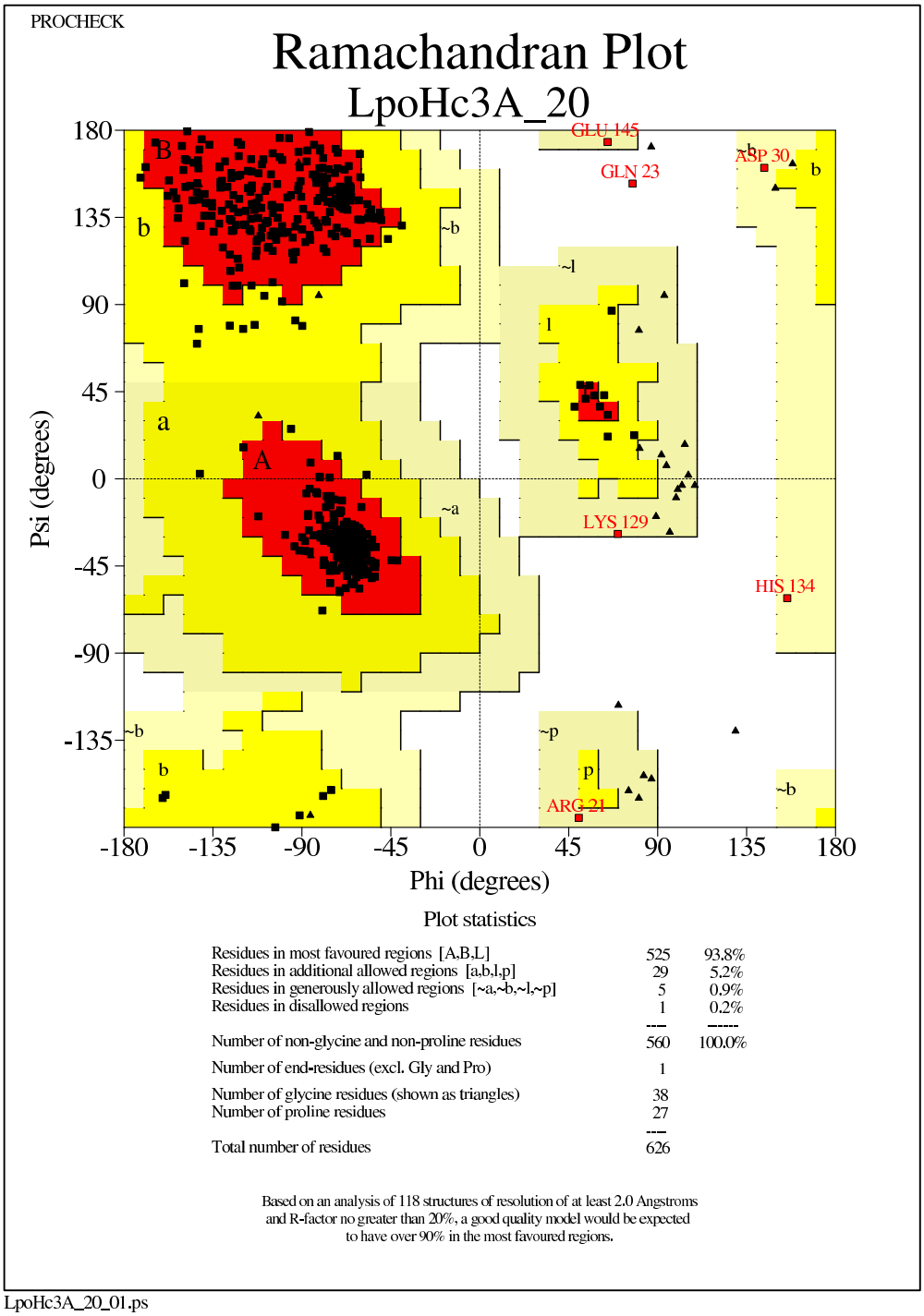


Figure 4.48: Ramachandran plot of the molecular model LpoHc2_20 representing *Limulus* subunit type II. Residues in 'disallowed' regions are $I^{25}(23)$, $D^{152}(150)$. White areas – disallowed regions; cream areas – generously allowed regions (∼a, ∼b, ∼l, ∼p); yellow areas – additional allowed regions (a, b, l, p); red areas – most favoured regions (A, B, L); triangles – glycine residues; black squares – non-glycine residues; red squares – residues in 'disallowed' regions



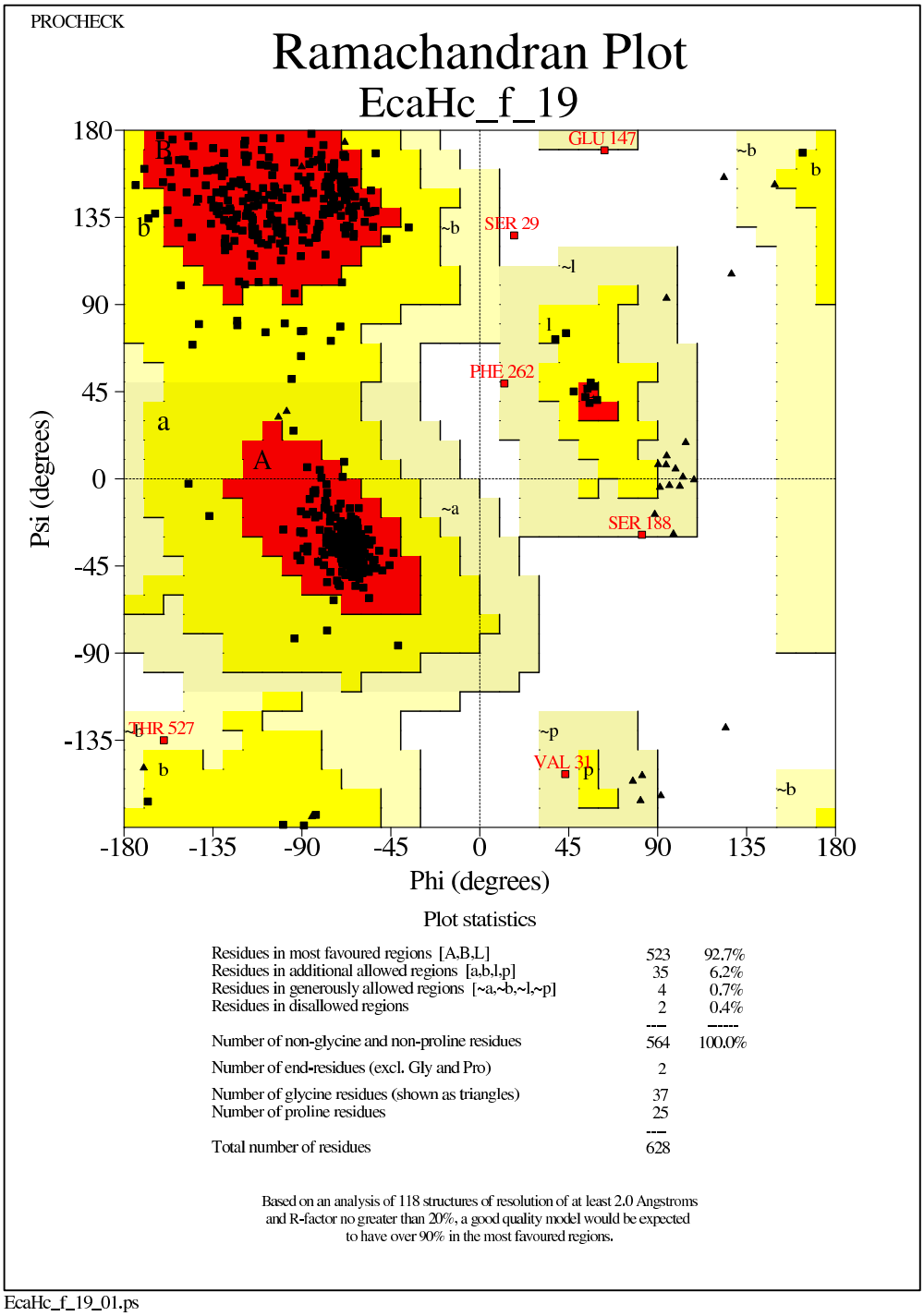
Plot statistics

Residues in most favoured regions [A,B,L]	525	93.8%
Residues in additional allowed regions [a,b,l,p]	29	5.2%
Residues in generously allowed regions [~a,~b,~l,~p]	5	0.9%
Residues in disallowed regions	1	0.2%
Number of non-glycine and non-proline residues	560	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	38	
Number of proline residues	27	
Total number of residues	626	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

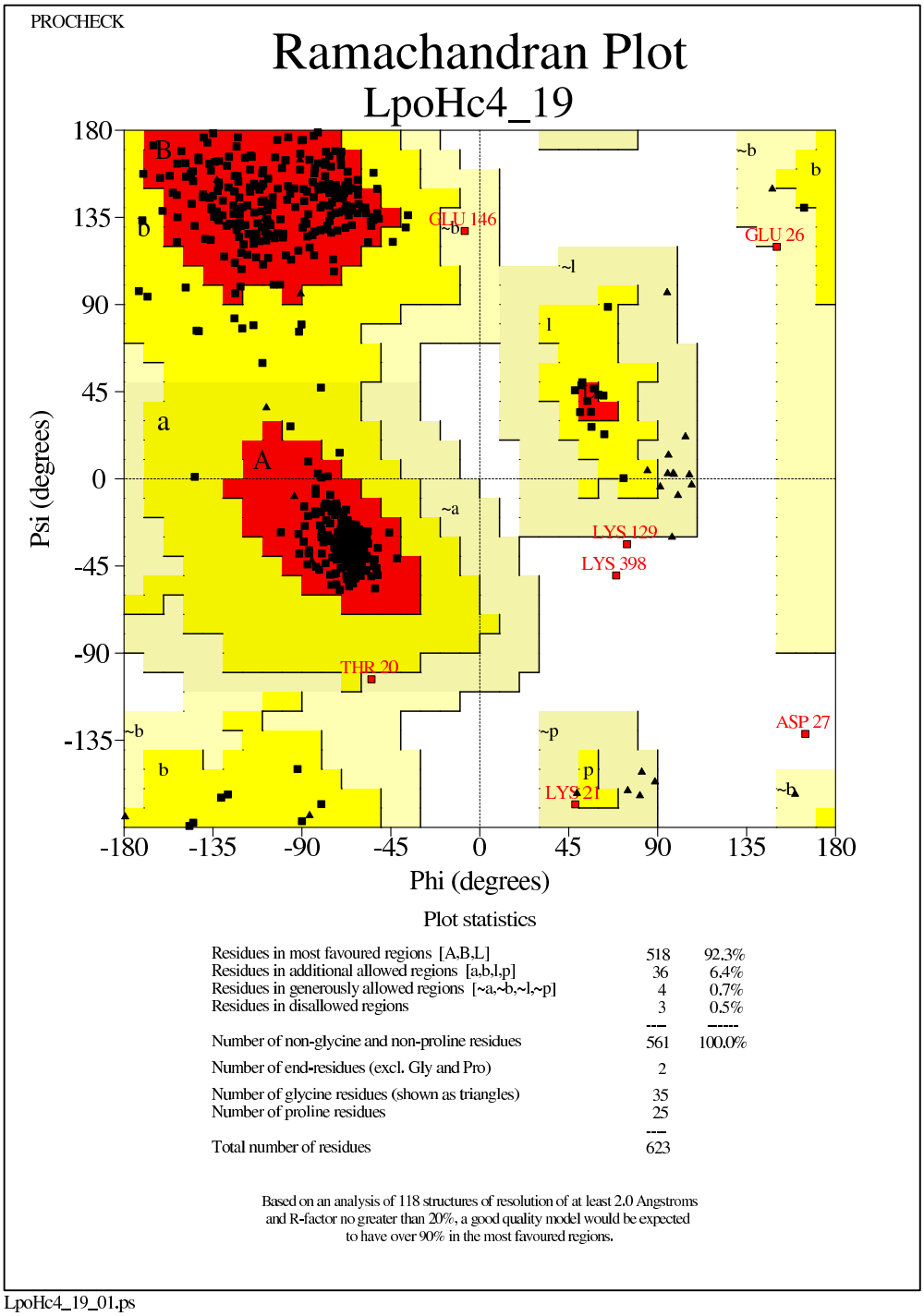
LpoHc3A_20_01.ps

Figure 4.49: Ramachandran plot of the molecular model LpoHc3A_20 representing *Limulus* subunit type IIIA. Residues in 'disallowed' regions are $R^{24}(21)$, $Q^{27}(23)$, $D^{34}(30)$, $K^{133}(129)$, $H^{138}(134)$, $E^{152}(145)$. White areas – disallowed regions; cream areas – generously allowed regions (~a, ~b, ~l, ~p); yellow areas – additional allowed regions (a, b, l, p); red areas – most favoured regions (A, B, L); triangles – glycine residues; black squares – non-glycine residues; red squares – residues in 'disallowed' regions



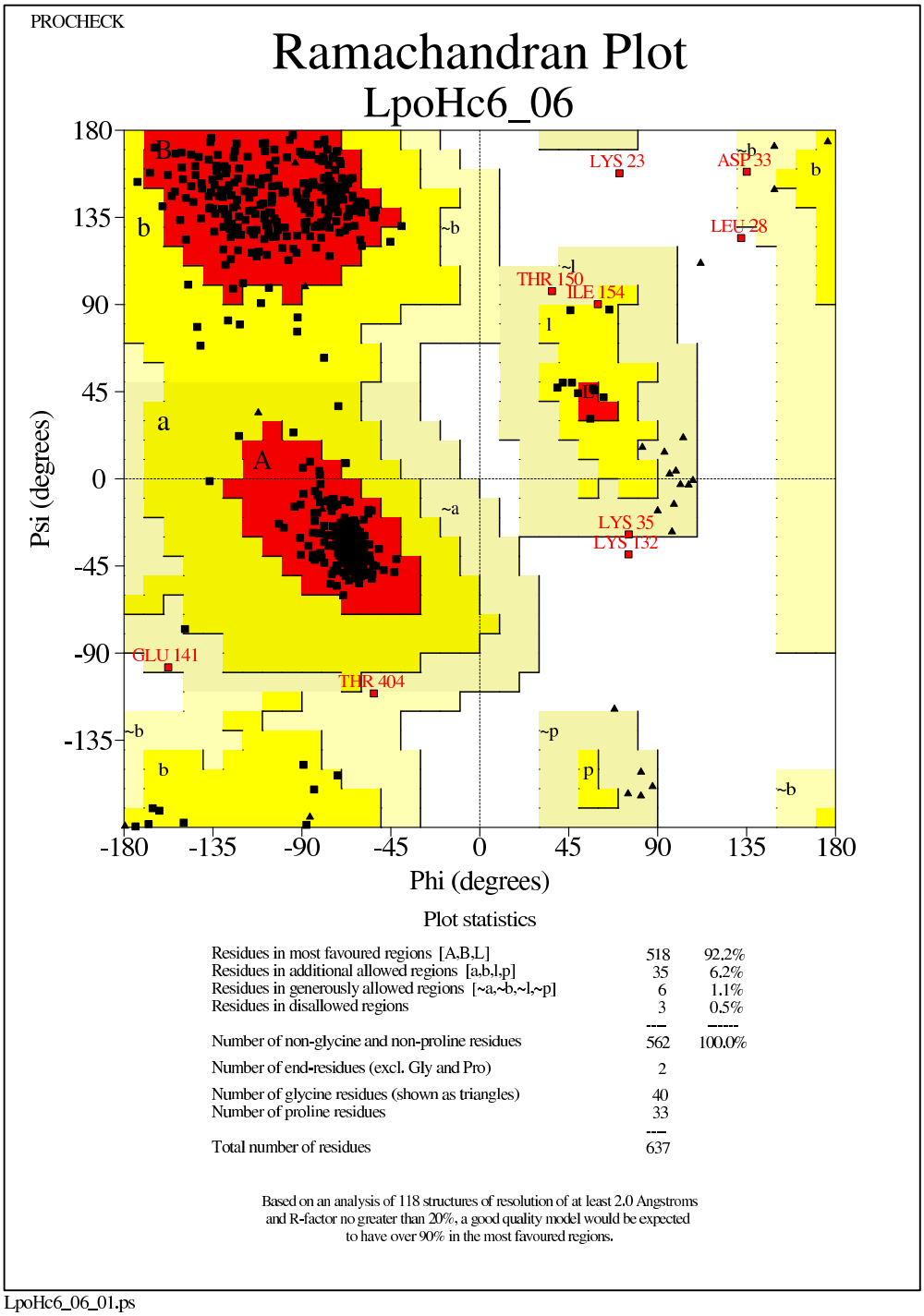
EcaHc_f_19_01.ps

Figure 4.50: Ramachandran plot of the molecular model EcaHc_f_19 representing *Limulus* subunit type IIIB. Residues in 'disallowed' regions are $S^{33}(29)$, $V^{35}(31)$, $E^{152}(147)$, $S^{193}(188)$, $F^{267}(262)$, $T^{534}(527)$. White areas – disallowed regions; cream areas – generously allowed regions (~a, ~b, ~l, ~p); yellow areas – additional allowed regions (a, b, l, p); red areas – most favoured regions (A, B, L); triangles – glycine residues; black squares – non-glycine residues; red squares – residues in 'disallowed' regions



LpoHc4_19_01.ps

Figure 4.51: Ramachandran plot of the molecular model LpoHc4_19 representing *Limulus* subunit type IV. Residues in 'disallowed' regions are $T^{23}(20)$, $K^{24}(21)$, $E^{30}(26)$, $D^{31}(27)$, $K^{133}(129)$, $E^{152}(146)$, $K^{404}(398)$. White areas – disallowed regions; cream areas – generously allowed regions (∼a, ∼b, ∼l, ∼p); yellow areas – additional allowed regions (a, b, l, p); red areas – most favoured regions (A, B, L); triangles – glycine residues; black squares – non-glycine residues; red squares – residues in 'disallowed' regions



LpoHc6_06_01.ps

Figure 4.52: Ramachandran plot of the molecular model LpoHc6_06 representing *Limulus* subunit types V and VI. Residues in 'disallowed' regions are $K^{24}(23)$, $L^{29}(28)$, $D^{34}(33)$, $K^{36}(35)$, $K^{133}(132)$, $E^{142}(141)$, $T^{152}(150)$, $I^{156}(154)$, $T^{407}(404)$. White areas – disallowed regions; cream areas – generously allowed regions (~a, ~b, ~l, ~p); yellow areas – additional allowed regions (a, b, l, p); red areas – most favoured regions (A, B, L); triangles – glycine residues; black squares – non-glycine residues; red squares – residues in 'disallowed' regions

Chapter 5

Conclusions and Outlook

Hexamerin The presented cryo-EM structure is the first median to high resolved structure of this class of proteins. The result confirmed that the quaternary structure observed in arthropod hemocyanins is applicable to the hexamerin LSP-2 from *Drosophila melanogaster* and most probably to all known insect storage proteins that belong to the hemocyanin superfamily.

Cryo-EM 3D-reconstruction combined with molecular modelling and rigid-body fitting showed a high correlation between additional structural features and amino acid insertions of the hexamerin LSP-2 when compared to the quaternary and primary structure of hemocyanin. This indicates that the results are highly significant.

Due to its small size and clear symmetry, the *Drosophila* hexamerin LSP-2 showed to be a good model to test new computational procedures for the cryo-EM single particle analysis. To optimize the molecular model, the technique of flexible fitting can be applied. This technique arranges a preliminary model calculated by satisfaction of spatial restraints into a cryo-EM structure, thus yielding an improved model by taking real structural features into account. Know-how gained on this 'simple' structure can then be transferred to higher-ordered oligomers.

Hemocyanin In the present cryo-EM structure of *Limulus polyphemus* hemocyanin, eleven types of inter-hexamer interfaces have been identified that in the native 8×6mer sum up to 46 inter-hexamer bridges: 24 within the four 2×6mers, 10 to establish the two 4×6mers, and 12 to assemble the two 4×6mers into an 8×6mer. From advanced molecular modelling and rigid-body fitting, each of these interfaces has been assigned to very few amino acids that now serve as candidates for the chemical bonds between the eight hexamers. In chelicerate hemocyanins, all hierarchical levels of the protein are involved to establish the oxygen binding properties, and reveal a stepwise decrease of cooperativity from 4×6mers to 2×6mers to 1×6mers ([Hartmann and Decker, 2002], [Savel-Niemann *et al.*, 1988], [Brouwer and Serigstad, 1989]). Consequently, allosteric interaction has to be somehow transferred via the inter-hexamer bridges.

Indeed, several intriguing structures have been unravelled as possible mediators of allosteric signals between the different levels of the nested hierarchy: the II – II interface, the II – IV interface and the V – VI interface between the 1×6mers, the VI – IIIB/IV/V interface between the 2×6mers, the IIIA – IIIA interface and the IV – IV interface between the 4×6mers. The present *Limulus* hemocyanin preparation has been studied under oxygen saturation conditions, in the presence of 25% oxygen, to produce vitrification of the protein molecules in their fully oxygenated state (under physiological conditions usually the R-state). Indeed, the retrieved parameters of the *Limulus* 4×6mer convincingly fit the SAXS model of the oxy–4×6mer of Eurypelma ([Hartmann and Decker, 2002]), whereas the deoxy–4×6mer SAXS model is significantly different (see Table 4.1).

I am optimistic that the same procedure, carried out under oxygen-free conditions (i.e., in 100% nitrogen), will produce the hemocyanin molecule in its fully deoxygenated state (usually the T-state). Molecular modelling with the structure of *Panulirus interruptus* hemocyanin as template, which is assumingly the deoxygenated R-state, may yield significant differences mainly in the region of domain 1. cryo-EM structure and analysis at a comparable resolution should then show the different inter-hexamer bridges in the deoxygenated 8×6mer, which might assist the understanding as to how, during oxygen binding, allosteric signals are structurally transferred between the different levels of the oligomeric hierarchy. Thus, the present cryo-EM structure opens the door to a fundamental understanding of the function of this highly cooperative protein.

Chapter 6

Summary

In the present study, the quaternary structures of *Drosophila melanogaster* hexamerin LSP-2 and *Limulus polyphemus* hemocyanin, both proteins from the hemocyanin superfamily, were elucidated to a 10 Å resolution with the technique of cryo-EM 3D-reconstruction. Furthermore, molecular modelling and rigid-body fitting allowed a detailed insight into the cryo-EM structures at atomic level. The results are summarised as follows:

Hexamerin

- The cryo-EM structure of *Drosophila melanogaster* hexamerin LSP-2 is the first quaternary structure of a protein from the group of the insect storage proteins.
- The hexamerin LSP-2 is a hexamer of six bean-shaped subunits that occupy the corners of a trigonal antiprism, yielding a D3 (32) point-group symmetry.
- Molecular modelling and rigid-body fitting of the hexamerin LSP-2 sequence showed a significant correlation between amino acid inserts in the primary structure and additional masses of the cryo-EM structure that are not present in the published quaternary structures of chelicerate and crustacean hemocyanins.
- The cryo-EM structure of *Drosophila melanogaster* hexamerin LSP-2 confirms that the arthropod hexameric structure is applicable to insect storage proteins.

Hemocyanin

- The cryo-EM structure of the 8×6mer *Limulus polyphemus* hemocyanin is the highest resolved quaternary structure of an oligo-hexameric arthropod hemocyanin so far.
- The hemocyanin is build of 48 bean-shaped subunits which are arranged in eight hexamers, yielding an 8×6mer with a D2 (222) point-group symmetry. The 'basic building blocks' are four 2×6mers that form two 4×6mers in an anti-parallel manner, latter aggregate 'face-to-face' to the 8×6mer.
- The morphology of the 8×6mer was gauged and described very precisely on the basis of the cryo-EM structure.
- Based on earlier topology studies of the eight different subunit types of *Limulus polyphemus* hemocyanin, eleven types of interhexamer interfaces have been identified that in the native 8×6mer sum up to 46 inter-hexamer bridges – 24 within the four 2×6mers, 10 to establish the two 4×6mers, and 12 to assemble the two 4×6mers into an 8×6mer.
- Molecular modelling and rigid-body fitting of *Limulus polyphemus* and orthologous *Erypelma californicum* sequences allowed to assign very few amino acids to each of these interfaces. These amino acids now serve as candidates for the chemical bonds between the eight hexamers.
- Most of the inter-hexamer contacts are conspicuously histidine-rich and evince constellations of amino acids that could constitute the basis for the allosteric interactions between the hexamers.
- The cryo-EM structure of *Limulus polyphemus* hemocyanin opens the door to a fundamental understanding of the function of this highly cooperative protein.

Chapter 7

Zusammenfassung

In der vorliegenden Arbeit wurden die Quartärstrukturen des Hexamerins LSP–2 von *Drosophila melanogaster* und des Hämocyanins von *Limulus polyphemus*, zwei Proteine der Hämocyanin-Superfamilie, mit einer Auflösung von 10 Å mittels Kryoelektronenmikroskopie (Kryo-EM) drei-dimensional rekonstruiert. Weiterhin wurde durch die Methoden des „molecular modelling“ und „rigid-body fitting“ eine detaillierte Beschreibung der Strukturen auf atomarem Niveau möglich. Die Ergebnisse lassen sich wie folgt zusammenfassen:

Hexamerin

- Die Kryo-EM-Struktur des Hexamerins LSP-2 von *Drosophila melanogaster* ist die erste hochaufgelöste Quartärstruktur eines Proteins der Gruppe der Insekten-Speicherproteine.
- Das Hexamerin LSP-2 ist ein Hexamer aus sechs nierenförmigen Untereinheiten, die die Ecken eines trigonalen Antiprismas besetzten und somit in einer D3 (32) Punktsymmetrie angeordnet sind.
- „Molecular modelling“ und „rigid-body fitting“ der Sequenz des Hexamerins LSP-2 ergaben eine klare Korrelation zwischen Insertionen von Aminosäuren in der Primärsstruktur und zusätzlichen Massen der Kryo-EM Struktur; letztere sind in den publizierten Quartärstrukturen der Cheliceraten- und Crustaceen-Hämocyanine nicht vorhanden.
- Die Kryo-EM-Struktur von LSP-2 bestätigt, dass die hexamere Struktur der Arthropoden-Hämocyanine auch bei den Insekten-Speicherproteinen vorliegt.

Hämocyanin

- Die Kryo-EM-Struktur des 8x6-Hämocyanins von *Limulus polyphemus* ist die bislang höchstaufgelöste Quartärstruktur eines oligo-hexameren Arthropoden-Hämocyanins.
- Das Hämocyanin ist aus 48 nierenförmige Untereinheiten aufgebaut, die in acht Hexameren (1×6mer) angeordnet sind, letztere nehmen eine D2 (222) Punktsymmetrie im nativen 8×6mer Molekül ein. Die strukturelle Grundeinheit bilden die vier 2×6mere, die in antiparalleler Ordnung zwei 4×6mere bilden; diese wiederum aggregieren „face-to-face“ zum 8×6mer.
- Die Morphologie des 8×6mers konnte auf Basis der Kryo-EM-Struktur erstmalig sehr genau vermessen und beschrieben werden.
- Basierend auf frühere topologische Studien der acht unterschiedlichen Untereinheitstypen des Hämocyanins von *Limulus polyphemus* wurden elf Arten von Verbindungen zwischen den Hexameren identifiziert, die im nativen Molekül eine Gesamtzahl von 46 Brücken ausbilden – 24 innerhalb der vier 2×6mere, zehn um beide 4×6mere zu bilden und letztlich 12 um diese zum 8×6mer zu verbinden.
- „Molecular modelling“ und „rigid-body fitting“ von *Limulus polyphemus* und orthologer *Erypelma californicum* Sequenzen ermöglichten die Identifizierung der Aminosäuren, welche die Brücken bilden. Viele dieser Aminosäuren sind überzeugende Kandidaten für chemische Bindungen zwischen den acht Hexameren.
- Die meisten der Inter-Hexamer-Kontakte sind auffällig histidinreich und weisen Aminosäurekonstellationen auf, die die Basis für die Übertragung allosterischer Wechselwirkungen zwischen den Hexameren darstellen könnten.
- Die Kryo-EM-Struktur des Hämocyanins von *Limulus polyphemus* öffnet so die Tür für ein grundlegendes Verständnis der Funktion dieses hochkooperativen Proteins.

Bibliography

[Adrian *et al.*, 1984] M. Adrian, J. Dubrochet, J. Lepault, and A. W. McDowall. Cryo-electron microscopy of viruses. *Nature*, 308:32–36, 1984.

[Averdam *et al.*, 2003] A. Averdam, J. Markl, and T. Burmester. Subunit sequences of the 4 x 6-mer hemocyanin from the golden orb-web spider, *Nephila inaurata*. *Eur. J. Biochem.*, 270(16):3432–9, 2003.

[Ballweber *et al.*, 2002] P. Ballweber, J Markl, and T. Burmester. Complete hemocyanin subunit sequences of the hunting spider *Cupiennius*, 2002.

[Beintema *et al.*, 1994] J. J. Beintema, W. T. Stam, B. Hazes, and M. P. Smidt. Evolution of arthropod hemocyanins and insect storage proteins (hexamerins). *Molecular Biology and Evolution*, 11(3):493–503, 1994.

[Benes *et al.*, 1990] H. Benes, R. G. Edmondson, P. Fink, J. Kejzlarova-Lepesant, J.-A. Lepesant, J. P. Miles, and D. W. Spivey. Adult expression to both the larval and adult fat body. *Dev. Biol.*, 142:138–142, 1990.

[Beresford *et al.*, 1997] P. J. Beresford, J. M. Basinski-Gray, J. K. Chiu, J. S. Chadwick, and W. P. Aston. Characterization of hemolytic and cytotoxic gallysins: a relationship with arylphorins. *Dev. Comp. Immunol.*, 21(3):253–66, 1997.

[Beverley and Wilson, 1982] S. M. Beverley and A. C. Wilson. Molecular evolution of *Drosophila* and higher diptera. I. Micro-complement fixation studies of a larval hemolymph protein. *J. Mol. Evol.*, 18:251–264, 1982.

[Beverley and Wilson, 1984] S. M. Beverley and A. C. Wilson. Molecular evolution in *Drosophila* and higher diptera. II. A time scale for fly evolution. *J. Mol. Evol.*, 21:1–13, 1984.

[Bijlholt and van Bruggen, 1986] M. Bijlholt and E. F. van Bruggen. A model for the architecture of the hemocyanin from the arthropod *Squilla mantis* (Crustacea, Stomatopoda). *Eur. J. Biochem.*, 155(2):339–44, 1986.

[Bijlholt *et al.*, 1979] M. M. Bijlholt, E. F. van Bruggen, and J. Bonaventura. Dissociation and reassembly of *Limulus polyphemus* hemocyanin. *Eur. J. Biochem.*, 95(2):399–405, 1979.

[Bijlholt *et al.*, 1982] M. M. Bijlholt, M. G. van Heel, and E. F. van Bruggen. Comparison of 4 x 6-meric hemocyanins from three different arthropods using computer alignment and correspondence analysis. *J. Mol. Biol.*, 161(1):139–53, 1982.

[Boisset *et al.*, 1990] N. Boisset, J. C. Taveau, and J. N. Lamy. An approach to the architecture of *Scutigera coleoptrata* hemocyanin by electron microscopy and image processing. *Biol. Cell*, 86:73–84, 1990.

[Boisset *et al.*, 1995] N. Boisset, P. Penczek, J. C. Taveau, J. Lamy, and J. Frank. Three-dimensional reconstruction of *Androctonus australis* hemocyanin labeled with a monoclonal fab fragment. *J. Struct. Biol.*, 115(1):16–29, 1995.

[Borland and van Heel, 1990] L. Borland and M. van Heel. Classification of image data in conjugate representation spaces. *J. opt. Soc. Am.*, A7:601–610, 1990.

[Böttcher *et al.*, 1997] B. Böttcher, S. A. Wynne, and R. A. Crowther. Determination of the fold of the core protein of hepatitis b virus by electron cryomicroscopy. *Nature*, 386:88–91, 1997.

[Braun and Wyatt, 1996] R. P. Braun and G. R. Wyatt. Sequence of the hexameric juvenile hormone-binding protein from the hemolymph of *Locusta migratoria*. *J. Biol. Chem.*, 271(49):31756–62, 1996.

[Brenowitz *et al.*, 1981] M. Brenowitz, C. Bonaventura, J. Bonaventura, and E. Gianazza. Subunit composition of high molecular weight oligomer: *Limulus polyphemus* hemocyanin. *Arch. Biochem. Biophys.*, 210(2):748–61, 1981.

[Brenowitz *et al.*, 1983] M. Brenowitz, C. Bonaventura, and J. Bonaventura. Assembly and calcium-induced cooperativity of *Limulus* IV hemocyanin: a model system for analysis of structure-function relationships in the absence of subunit heterogeneity. *Biochemistry*, 22(20):20, 1983.

[Brenowitz *et al.*, 1984] M. Brenowitz, C. Bonaventura, and J. Bonaventura. Self-association and oxygen-binding characteristics of the isolated subunits of *Limulus polyphemus* hemocyanin. *Arch. Biochem. Biophys.*, 230:238–249, 1984.

[Brock and Roberts, 1983] H. W. Brock and D. B. Roberts. An immunological and electrophoretic study of the larval serum proteins of *Drosophila* species. *Insect Biochem.*, 13:57–63, 1983.

[Brouwer and Serigstad, 1989] M. Brouwer and B. Serigstad. Allosteric control in *Limulus polyphemus* hemocyanin: functional relevance of interactions between hexamers. *Biochemistry*, 28(22):8819–27, 1989.

[Brouwer *et al.*, 1977] M. Brouwer, C. Bonaventura, and J. Bonaventura. Oxygen binding by *Limulus polyphemus* hemocyanin: allosteric modulation by chloride ions. *Biochemistry*, 16(17):3897–902, 1977.

[Burmester and Scheller, 1996] T. Burmester and K. Scheller. Common origin of arthropod tyrosinase, arthropod hemocyanin, insect hexamerin, and dipteran arylphorin receptor. *Journal of Molecular Evolution*, 42(6):713–728, 1996.

[Burmester *et al.*, 1998] T. Burmester, H. C. Massey, S. O. Zakharkin, and H. Benes. The evolution of hexamerins and the phylogeny of insects. *Journal of Molecular Evolution*, 47(1):93–108, 1998.

[Burmester, 1999] T. Burmester. Evolution and function of the insect hexamerins. *Eur. J. Entomol.*, 96(3):213–225, 1999.

[Burmester, 2001] T. Burmester. Molecular evolution of the arthropod hemocyanin superfamily. *Mol. Biol. Evol.*, 18(2):184–195, 2001.

[Burmester, 2002] Thorsten Burmester. Origin and evolution of arthropod hemocyanins and related proteins. *J. Comp. Physiol. [B]*, 172:95–117, 2002.

[Chacón and Wriggers, 2002] P. Chacón and W. Wriggers. Multi-resolution contour-based fitting of macromolecular structures. *J. Mol. Biol.*, 317(3):375–84, 2002.

[Chrysanthis *et al.*, 1994] G. Chrysanthis, A. D. Kaliafas, and A. C. Mintzas. Biosynthesis and tissue distribution of four major larval serum proteins during development of *Ceratitis capitata* (diptera). *Insect Biochemistry and Molecular Biology*, 24(8):811–818, 1994.

[de Haas and van Bruggen, 1994] F. de Haas and E. F. J. van Bruggen. The inter-hexameric contacts in the 4-hexameric hemocyanin from the tarantula *Eurypelma californicum* - a tentative mechanism for cooperative behavior. *Journal of Molecular Biology*, 237(4):464–478, 1994.

[Decker *et al.*, 1989] H. Decker, A. Savel-Niemann, D. Körschenhausen, E. Eckerson, and J. Markl. Allosteric oxygen-binding properties of reassembeled tarantula (*Eurypelma californicum*) hemocyanin with incorporated apo- or met-subunits. *Biol. Chem. HS*, 370:511–523, 1989.

[DeRosier and Klug, 1968] D. J. DeRosier and A. Klug. Reconstruction of three-dimensional structures from electron micrographs. *Nature*, 217:130–134, 1968.

[Dierkensen *et al.*, 1992] K. Dierkensen, D. Typke, R. Hegerl, A. J. Koster, and W. Baumeister. Towards atomic electron tomography. *Ultramicroscopy*, 40:71–87, 1992.

[Dolashka-Angelova *et al.*, 2005] P. Dolashka-Angelova, A. Dolashki, S. Stevanovic, R. Hristova, B. Atanasov, P. Nikolov, and W. Voelter. Structure and stability of arthropodan hemocyanin *Limulus polyphemus*. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 61(6):1207–17, 2005.

[Dube *et al.*, 1993] P. Dube, P. Tavares, R. Lurz, and M. van Heel. The portal protein of bacteriophage SPP1: A DNA pump with 13-fold symmetry. *EMBO J.*, 12(4):1303–1309, 1993.

[Ellerton *et al.*, 1983] H. D. Ellerton, N. F. Ellerton, and H. A. Robinson. Hemocyanin – a current perspective. *Proc. Biophys. Mol. Biol.*, 41:143–248, 1983.

[Gaykema *et al.*, 1984] W. P. Gaykema, W. G. Hol, J. M. Vereijken, N. M. Soeter, H. J. Bak, and J. J. Beintema. 3.2 Å structure of the copper-containing, oxygen-carrying protein *Panulirus interruptus* heamocyanin. *Nature*, 309(3):23–29, 1984.

[Gaykema *et al.*, 1986] W. P. Gaykema, A. Volbeda, and W. G. Hol. Structure determination of *Panulirus interruptus* haemocyanin at 3.2 Å resolution. Successful phase extension by sixfold density averaging. *J. Mol. Biol.*, 187(2):255–75, 1986.

[Harauz and Ottensmeyer, 1984] G. Harauz and F. P. Ottensmeyer. Direct three-dimensional reconstruction for macromolecular complexes from electron micrographs. *Ultramicroscopy*, 12:309–319, 1984.

[Harauz and van Heel, 1986a] G. Harauz and M. van Heel. *Direct 3D reconstruction from projections with initially unknown angles.* Pattern Recognition in Practice II. North-Holland Publishing, Amsterdam, 1986.

[Harauz and van Heel, 1986b] G. Harauz and M. van Heel. Exact filters for general geometry three dimensional reconstruction. *Optik*, 73:146–156, 1986.

[Harris, 1997] J. R. Harris. *Negative staining and cryoelectron microscopy.*, volume 35 of *RMS Microscopy Handbook*. BIOS Scientific Publishers Ltd., Oxford, UK., 1997.

[Harris, 1999] J. R. Harris. Negative staining of thinly spread biological particulates. *Methods in Molecular Biology*, 117:13–29, 1999.

[Hartmann and Decker, 2002] H. Hartmann and H. Decker. All hierarchical levels are involved in conformational transitions of the 4×6-meric tarantula hemocyanin upon oxygenation. *Biochim. Biophys. Acta*, 1601(2):132–7, 2002.

[Hazes *et al.*, 1993] B. Hazes, K. A. Magnus, C. Bonaventura, J. Bonaventura, Z. Dauter, K. H. Kalk, and W. G. Hol. Crystal structure of deoxygenated *Limulus polyphemus* subunit II hemocyanin at 2.18 Å resolution: clues for a mechanism for allosteric regulation. *Protein Sci.*, 2(4):597–619, 1993.

[Higgins and Sharp, 1988] D. G. Higgins and P. M. Sharp. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, 73:237–244, 1988.

[Hoppe *et al.*, 1974] W. Hoppe, J. Gaßmann, N. Hunsmann, H. J. Schramm, and M. Sturm. Three-dimensional reconstruction of individual negatively stained yeast fatty acid synthetase molecules from tilt-series in the electron microscope. *Hoppe-Seyler's Z. Physiol. Chem.*, 355:1483–1487, 1974.

[Ismail and Gillot, 1995] S. M. Ismail and C. Gillot. Identification, characterization, and developmental profile of a high molecular weight, juvenile hormone-binding protein in the hemolymph of the migratory grasshopper, *Melanoplus sanguinipes*. *Archives of Insect Biochemistry and Physiology*, 29(4):415–430, 1995.

[Jung *et al.*, 2005] H. Jung, Y. H. Kim, and S. Kim. Structural basis for the presence of a monoglucosylated oligosaccharide in mature glycoproteins. *Biochemical and Biophysical Research Communications*, 331(1):100–106, 2005.

[Karlsen *et al.*, 2005] O. A. Karlsen, L. Ramsevik, L. J. Bruseth, Ø. Larsen, A. Brenner, F. S. Berven, H. B. Jensen, and J. R. Lillehaug. Characterization of a prokariotic haemerythrin from the methanotrophic bacterium *Methylococcus capsulatus* (Bath). *FEBS Journal*, 272:2428–2440, 2005.

[Klaholz *et al.*, 2004] B. P. Klaholz, A. G. Myasnikov, and M. van Heel. Visualization of release factor 3 on the ribosome during termination of protein synthesis. *Nature*, 427:862–865, 2004.

[Lamy *et al.*, 1977] J. Lamy, J. Lamy, M. C. Baglin, and J. Weill. *Scorpion hemocyanin subunits: properties, dissociation, association.* Structure and function of haemocyanin. Springer, Berlin, Heidelberg, New York, 1977.

[Lamy *et al.*, 1979] J. Lamy, J. Lamy, J. Weill, J. Bonaventura, C. Bonaventura, and M. Brenowitz. Immunological correlates between the multiple hemocyanin subunits of *Limulus polyphemus* and *Tachypleus tridentatus*. *Arch. Biochem. Biophys.*, 196:324–339, 1979.

[Lamy *et al.*, 1981] J. Lamy, M. C. Bijlholt, P. Y. Sizaret, and E. F. van Bruggen. Quaternary structure of scorpion (*androctonus australis*) hemocyanin. Localization of subunits with immunological methods and electron microscopy. *Biochemistry*, 20(7):1849–56, 1981.

[Lamy *et al.*, 1982] J. Lamy, P. Y. Sizaret, J. Frank, A. Verschoor, R. Feldmann, and J. Bonaventura. Architecture of *Limulus polyphemus* hemocyanin. *Biochemistry*, 21(26):6825–6833, 1982.

[Lamy *et al.*, 1983] J. Lamy, J. Lamy, P. Y. Sizaret, P. Billiald, P. Jolles, J. Jolles, R. J. Feldmann, and J. Bonaventura. Quaternary structure of *Limulus polyphemus* hemocyanin. *Biochemistry*, 22(24):5573–5583, 1983.

[Lamy *et al.*, 1996] J. N. Lamy, B. N. Green, A. Toulmond, J. S. Wall, R. E. Weber, and S. N. Vinogradov. Giant hexagonal bilayer hemoglobins. *Chemical Reviews*, 96(8):3113–3124, 1996.

[Lang and van Holde, 1991] W. H. Lang and K. E. van Holde. Cloning and sequencing of *Octopus dofleini* hemocyanin cDNA: derived sequences of functional units Ode and Odf. *Proc. Natl. Acad. Sci. U S A*, 88(1):244–8, 1991.

[Laskowski *et al.*, 1993] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26:283–291, 1993.

[Lebart *et al.*, 1984] L. Lebart, A. Morineau, and K. M. Warwick. *Multivariate descriptive statistical analysis*. Wiley, New York, 1984.

[Lieb *et al.*, 2006] B. Lieb, K. Dimitrova, H.-S. Kang, S. Braun, W. Gebauer, A. Martin, B. Hanelt, S. A. Saenz, C. M. Adema, and Markl J. Red blood with blue-blood ancestry: intriguing structure of a snail hemoglobin. *Proc. Natl. Acad. Sci.*, 103(32):12011–6, 2006.

[Loewe, 1978] R. Loewe. Hemocyanins in spiders. V. Fluorometric recording of oxygen binding curves, and its application to the analysis of allosteric interactions in *Eurypelma californicum* hemocyanin. *J. Comp. Physiol. B.*, 128:161–168, 1978.

[Ludtke *et al.*, 1999] S. J. Ludtke, P. R. Baldwin, and W. Chiu. EMAN: Semiautomated software for high-resolution single- particle reconstructions. *Journal of Structural Biology*, 128(1):82–97, 1999.

[Ludtke *et al.*, 2004] S. J. Ludtke, D. H. Chen, J. L. Song, D. T. Chuang, and W. Chiu. Seeing GroEL at 6 Å resolution by single particle electron cryomicroscopy. *Structure*, 12(7):1129–1136, 2004.

[Magee *et al.*, 1994] J. Magee, N. Kraynack, H. C. Massey, and W. H. Telfer. Properties and significance of a riboflavin-binding hexamerin in the hemolymph of *Hyalophora cecropia*. *Arch. Insect. Biochem. Physiol.*, 25(2):137–57, 1994.

[Magnus *et al.*, 1991] K. A. Magnus, E. E. Lattman, A. Volbeda, and W. G. Hol. Hexamers of subunit II from *Limulus* hemocyanin (a 48-mer) have the same quaternary structure as whole *Panulirus* hemocyanin molecules. *Proteins*, 9(4):240–7, 1991.

[Magnus *et al.*, 1994] K. A. Magnus, B. Hazes, H. Ton That, C. Bonaventura, J. Bonaventura, and W. G. Hol. Crystallographic analysis of oxygenated and deoxygenated states of arthropod hemocyanin shows unusual differences. *Proteins*, 19(4):302–9, 1994.

[Mangum *et al.*, 1985] C. P. Mangum, J. L. Scott, R. E. L. Black, K. I. Miller, and K. E. van Holde. Centipedal hemocyanin: its structure and its implications for arthropod phylogeny. *Proc. Natl. Acad. Sci. USA*, 82:3721–3725, 1985.

[Markl and Decker, 1992] J. Markl and H. Decker. Molecular structure of the arthropod hemocyanins. In CH. P. Mangum, editor, *Blood and Tissue Oxygen Carriers*, volume 13 of *Advances in Comparative and Environmental Physiology*, pages 325–376. Springer Verlag, Berlin, 1992.

[Markl and Winter, 1989] J. Markl and S. Winter. Subunit-specific monoclonal antibodies to tarantula hemocyanin, and a common epitope shared with calliphorin. *J. Comp. Physiol. B.*, 159:139–151, 1989.

[Markl *et al.*, 1982] J. Markl, H. Decker, B. Linzen, W. G. Schutter, and E. F. J. van Bruggen. Hemocyanins in spiders. XV. the role of individual subunits in the assembly of *Eurypelma* hemocyanin. *Hoppe Seyler's Z. Physiol. Chem.*, 363:73–87, 1982.

[Markl *et al.*, 1992] J. Markl, T. Burmester, H. Decker, A. Savel-Niemann, J. R. Harris, M. Suling, U. Naumann, and K. Scheller. Quaternary and subunit structure of *Calliphora* arylphorin as deduced from electron microscopy, electrophoresis, and sequence similarities with arthropod hemocyanin. *J. Comp. Physiol. [B]*, 162(8):665–80, 1992.

[Markl, 1980] J. Markl. Hemocyanins in spiders. XI. The quaternary structure of *Cupiennius* hemocyanin. *J. Comp. Physiol. B.*, 140:199–207, 1980.

[Markl, 1986] J. Markl. Evolution and function of structurally diverse subunits in the respiratory protein hemocyanin from arthropods. *Biol. Bull.*, 171:190–115, 1986.

[Meissner *et al.*, 2003] U. Meissner, M. Stohr, K. Kusche, T. Burmester, H. Stark, J. R. Harris, E. V. Orlova, and J. Markl. Quaternary structure of the european spiny lobster (*Palinurus elephas*) 1 × 6-mer hemocyanin from cryoEM and amino acid sequence data. *J. Mol. Biol.*, 325:99–109, 2003.

[Meissner *et al.*, 2004] U. Meissner, A. G. Martin, B. O. Schwarz, M. Stohr, W. Gebauer, Harris J. R., and Markl J. 3-D reconstruction of hemocyanins and other invertebrate hemolymph proteins by cryo-TEM: an overview. *Micron*, 35(1-2):7–9, 2004.

[Meissner *et al.*, 2007] U. Meissner, E. Schroder, D. Scheffler, A. G. Martin, and J. R. Harris. Formation, TEM study and 3D reconstruction of the human erythrocyte peroxiredoxin-2 dodecahedral higher-order assembly. *Micron*, 38(1):29–39, 2007.

[Mindell and Grigorieff, 2003] J. A. Mindell and N. Grigorieff. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology*, 142:334–347, 2003.

[Miura *et al.*, 1994] K. Miura, M. Nakagawa, Y. Chinzei, T. Shinoda, E. Nagao, and H. Numata. Structural and functional studies on biliverdi-associated cyanoprotein from the bean bug, *Riptortus clavatus*. *Zool. Sci.*, 11:537–545, 1994.

[Mousseron-Grall *et al.*, 1997] S. Mousseron-Grall, J. Kejzlarova-Lepesant, T. Burmester, C. Chihara, M. Barray, E. Delain, R. Pictet, and J. A. Lepesant. Sequence, structure and evolution of the ecdysone-inducible Lsp-2 gene of *Drosophila melanogaster*. *European Journal of Biochemistry*, 245(1):191–198, 1997.

[Munn and Greville, 1969] E. A. Munn and G. D. Greville. The soluble proteins of developing *Calliphora erythrocephala*, particularly calliphorin, and similar proteins in other insects. *J. Insect. Physiol.*, 15:1935–50, 1969.

[Orlova and van Heel, 1994] E. Orlova and M. van Heel. Angular reconstitution of macromolecules with arbitrary point group symmetry. *Proceedings ICEM 13 – Paris 1994*, 1:507–508, 1994.

[Pallavicini *et al.*, 2001] A. Pallavicini, E. Negrisolo, R. Barbato, S. Dewilde, A. Ghiretti-Magaldi, L. Moens, and G. Lanfranchi. The primary structure of globin and linker chains from the chlorocruorin of the polychaete *Sabella spallanzanii*. *J. Biol. Chem.*, 276(28):26384–26390, 2001.

[Peter and Scheller, 1991] M. G. Peter and K. Scheller. Arylphorins and the integument. *The Physiology of Insect Epidermis.*, CSIRO Publications, Melbourne, pp.:115–124, 1991.

[Pettersen *et al.*, 2004] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25(13):1605–1612, 2004.

[Phipps *et al.*, 1994] D. J. Phipps, J. S. Chadwick, and W. P. Aston. Gallysin-1, an antibacterial protein isolated from the hemolymph of *Galleria mellonella*. *Dev. Comp. Immunol.*, 18:13–23, 1994.

[Powell *et al.*, 1984] D. Powell, J. D. Sato, H. W. Brock, and D. B. Roberts. Regulation of synthesis in the larval serum proteins of *Drosophila melanogaster*. *Dev. Biol.*, 102:206–216, 1984.

[Radermacher *et al.*, 1987] M. Radermacher, T. Wagenknecht, A. Verschoor, and J. Frank. Three-dimensional structure of the large ribosomal subunit from *Escherischia coli*. *EMBO J.*, 6:1107–1114, 1987.

[Radermacher, 1988] M. Radermacher. Three-dimensional reconstruction of single particles from random and non-random tilt series. *J. elect. Microsc. Tech.*, 9:359–394, 1988.

[Radon, 1917] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte langs gewisser Manigfaltigkeiten. Berichte über die Verhandlungen der Königlich Sachsichen Gesellschaft der Wissenschaften zu Leipzig. *Math. Phys. Klasse*, 69:262–277, 1917.

[Rawat *et al.*, 2006] U. Rawat, H. Gao, A. Zavialov, R. Gursky, M. Ehrenberg, and J. Frank. Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM. *J. Mol. Biol.*, 357(4):1144–53, 2006.

[Redmond and Gamlin, 1999] R. W. Redmond and J. N. Gamlin. A compilation of singlet oxygen yields from biologically relevant molecules. *Photochemistry and Photobiology*, 70:391–475, 1999.

[Roberts *et al.*, 1977] D. B. Roberts, J. Wolfe, and M. E. Akam. The developmental profiles of two major haemolymph proteins from *Drosophila melanogaster*. *J. Insect. Physiol.*, 23(7):871–8, 1977.

[Roberts *et al.*, 1991a] D. B. Roberts, T. Jowett, J. Hughes, D. F. Smith, and D. M. Glover. The major serum protein of *Drosophila* larvae, larval serum protein 1, is dispensable. *Eur. J. Biochem.*, 195(1):195–201, 1991.

[Roberts *et al.*, 1991b] D. B. Roberts, J. D. Turing, and S. A. R. Loughlin. The advantages that accrue to *Drosophila melanogaster* possessing larval serum protein 1. *J. Insect. Physiol.*, 37(5):391–400, 1991.

[Roberts, 1987] D. B. Roberts. The functions of the major serum proteins of *Drosophila* larvae. *Biol. Chem. Hoppe-Seyler*, 368:572, 1987.

[Sali and Blundell, 1993] A. Sali and T. L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, 234:779–815, 1993.

[Savel-Niemann *et al.*, 1988] A. Savel-Niemann, J. Markl, and B. Linzen. Hemocyanins in spiders XXII. range of allosteric interaction in a four-hexamer hemocyanin. Co-operativity and bohr effect in dissociation intermediates. *J. Mol. Biol.*, 204:385–395., 1988.

[Schatz *et al.*, 1995] M. Schatz, E. V. Orlova, P. Dube, J. Jager, and M. Vanheel. Structure of *Lumbricus terrestris* hemoglobin at 30 Å resolution determined using angular reconstitution. *Journal of Structural Biology*, 114(1):28–40, 1995.

[Serysheva *et al.*, 1995] I. I. Serysheva, E. V. Orlova, W. Chiu, M. B. Sherman, S. L. Hamilton, and M. van Heel. Electron cryomicroscopy and angular reconstitution

used to visualize the skeletal-muscle calcium-release channel. *Nature Structural Biology*, 2(1):18–24, 1995.

[Shirras and Bownes, 1989] A. D. Shirras and M. Bownes. Cricklet: a locus regulating a number of adult functions of *Drososphila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 86:4559–4563, 1989.

[Stark *et al.*, 1996] H. Stark, F. Zemlin, and C. Boettcher. Electron radiation damage to protein crystals of bacteriorhodopsin at different temperatures. *Ultramicroscopy*, 63:75–79, 1996.

[Stenkamp, 1994] R. E. Stenkamp. Dioxygen and hemerythrin. *Chem. Rev.*, 94:715–726, 1994.

[Sullivan *et al.*, 1974] B. Sullivan, J. Bonaventura, and C. Bonaventura. Functional differences in the multiple hemocyanins of the horseshoe crab, *Limulus polyphemus*. *Proc. Natl. Acad. Sci. USA*, 71:2558–2562, 1974.

[Taveau *et al.*, 1997] J. C. Taveau, N. Boisset, J. Lamy, O. Lambert, and J. N. Lamy. Three-dimensional reconstruction of *Limulus polyphemus* hemocyanin from cryoelectron microscopy. *J. Mol. Biol.*, 266(5):1002–15, 1997.

[Telfer and Kunkel, 1991] W. H. Telfer and J. G. Kunkel. The function and evolution of insect storage hexamers. *Annu. Rev. Entomol.*, 36:205–228, 1991.

[Telfer *et al.*, 1983] W. H. Telfer, P. S. Keim, and J. H. Law. Arylphorin, a new protein from *Hyalophora cecropia*: Comparisons with calliphorin and manducin. *Insect Biochem.*, 13(6):601–13, 1983.

[Terwilliger *et al.*, 1999] N.B. Terwilliger, L. Dangott, and M. Ryan. Cryptocyanin, a crustacean molting protein: evolutionary link with arthropod hemocyanins and insect hexamerins. *Proc. Natl. Acad. Sci.*, 96(5):2013–8, 1999.

[Terwilliger, 1982] N. B. Terwilliger. Effect of subunit composition on the quaternary structure of isopod (*Ligia pallasii*) hemocyanin. *Biochemistry*, 21:2579–2586, 1982.

[Thompson *et al.*, 1997] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 4:4876–4882, 1997.

[Thon, 1966] F. Thon. Zur Defokussierungsabhängigkeit des Phasenkontrastes bei der elektronenmikroskopischen Abbildung. *Z. Naturforschung*, 21a:476–478, 1966.

[Topham *et al.*, 1986] R. W. Topham, S. Tesh, C. Bonaventura, and J. Bonaventura. *Active-site heterogeneity as revealed by peroxide and mercury: interactions with purified subunits of Limulus hemocyanin.* Invertebrate oxygen carriers. Springer, Berlin, Heidelberg, New York, 1986.

[Tsakas *et al.*, 1991] S. Tsakas, P. G. Katsoris, K. Bourtzis, and V. J. Marmaras. Incorporation of arylphorins (LSP-1) and LSP-2 like protein into the integument of *Ceratitis capitata* during pupariation. *Insect Biochem.*, 21(5):507–15, 1991.

[Vagin and Teplyakov, 1997] A. Vagin and A. Teplyakov. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.*, 30:1022–1025, 1997.

[van Bruggen *et al.*, 1980] E. F. J. van Bruggen, M. M. C Bijlholt, W. G. Schutter, T. Wichertjes, J. Bonaventura, C. Bonaventura, J. Lamy, J. Lamy, M. Leclerc, H.-J. Schneider, J. Markl, and B. Linzen. The role of structurally diverse subunits in the assembly of three cheliceratan hemocyanins. *FEBS Letters*, 116(2):207–210, 1980.

[van Bruggen *et al.*, 1981] E. F. J. van Bruggen, W. G. Schutter, J. F. L. van Breemen, M. M. C. Bijlholt, and T. Wichertjes. Arthropodan and molluscan haemocyanins. *Electron Microscopy of Proteins*, 1:1–38, 1981.

[van Bruggen, 1983] E. F. J. van Bruggen. *An electron microscopists view of the quaternary structure of arthropodan and molluscan hemocyanins.* Structure and function of invertebrate respiratory proteins. Life Chem. Rep. Suppl. 1, Harwood, London, 1983.

[van Heel and Dube, 1994] M. van Heel and P. Dube. Quaternary structure of multihexameric arthropod hemocyanins. *Micron*, 25(4):387–418, 1994.

[van Heel and Frank, 1981] M. van Heel and J. Frank. Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, 6:187–194, 1981.

[van Heel and Keegstra, 1981] M. van Heel and W. Keegstra. Imagic: a fast, flexible and friendly image analysis software system. *Ultramicroscopy*, 7:113–130, 1981.

[van Heel and Schatz, 2005] M. van Heel and M. Schatz. Fourier shell correlation threshold criteria. *J. Struct. Biol.*, 151(3):250–62, 2005.

[van Heel and Stöffler-Meilicke, 1985] M. van Heel and M. Stöffler-Meilicke. Characteristic views of *E. coli* and *B. stearothermophilus* 30S ribosomal subunits in the electron microscope. *EMBO J*, 4(9):2389–95, 1985.

[van Heel *et al.*, 1992a] M. van Heel, M. Schatz, and E. Orlova. Correlation functions revisited. *Ultramicroscopy*, 46(1-4):307–316, 1992.

[van Heel *et al.*, 1992b] M. van Heel, H. Winkler, E. Orlova, and M. Schatz. Structure analysis of ice-embedded single particles. *Scanning Microsc. Suppl.*, 6:23–42, 1992.

[van Heel *et al.*, 1996] M. van Heel, G. Harauz, E. V. Orlova, R. Schmidt, and M. Schatz. A new generation of the imagic image processing system. *J. Struct. Biol.*, 116(1):17–24, 1996.

[van Heel *et al.*, 2000] M. van Heel, B. Gowen, R. Matadeen, E. V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, and A. Patwardhan. Single-particle electron cryo-microscopy: towards atomic resolution. *Q. Rev. Biophys.*, 33(4):307–369, 2000.

[van Heel, 1984a] M. van Heel. Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy*, 13:111–124, 1984.

[van Heel, 1984b] M. van Heel. Three-dimensional reconstructions from projections with unknown angular relationship. *Proc. 8th Eur. Cong. on EM, Budapest*, 2:1347–1348, 1984.

[van Heel, 1987] M. van Heel. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, 21:111–124, 1987.

[van Heel, 1989] M. van Heel. Classification of very large electron microscopal image data sets. *Optik*, 82:114–126, 1989.

[van Holde and Miller, 1982] K. E. van Holde and K. J. Miller. Haemocyanins. *Q. Rev. Biophys.*, 15:1–129, 1982.

[van Holde *et al.*, 1992] K. E. van Holde, K. I. Miller, and W. H. Lang. Molluscan hemocyanins: Structure and function. *Advanced in Comparative and Environmental Physiology*, 13:252–300, 1992.

[van Schaick *et al.*, 1982] E. J. van Schaick, W. G. Schutter, W. P. Gaykema, A. M. Schepman, and W. G. Hol. Structure of *Panulirus interruptus* hemocyanin at 5 Å resolution. *J. Mol. Biol.*, 158(3):457–85, 1982.

[Voit *et al.*, 2000] R. Voit, G. Feldmaier-Fuchs, T. Schweikardt, H. Decker, and T. Burmester. Complete sequence of the 24-mer hemocyanin of the tarantula *Eurypelma californicum* - Structure and intramolecular evolution of the subunits. *Journal of Biological Chemistry*, 275(50):39339–39344, 2000.

[Volbeda and Hol, 1989a] A. Volbeda and W. G. Hol. Crystal structure of hexameric haemocyanin from *Panulirus interruptus* refined at 3.2 Å resolution. *J. Mol. Biol.*, 209(2):249–79, 1989.

[Volbeda and Hol, 1989b] A. Volbeda and W. G. Hol. Pseudo 2-fold symmetry in the copper-binding domain of arthropodan haemocyanins. possible implications for the evolution of oxygen transport proteins. *J. Mol. Biol.*, 206(3):531–46, 1989.

[Ward, 1982] J. H. Jr. Ward. Hierarchical grouping to optimize an objective function. *Am. Statist. Assoc.*, 58:236–244, 1982.

[Wriggers *et al.*, 1999] W. Wriggers, R. A. Milligan, and J. A. McCammon. Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.*, 125:185–195, 1999.

[Zhu *et al.*, 1997] J. Zhu, P.A. Penczek, R. Schroder, and J. Frank. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S *Escherichia coli* ribosome. *Journal of Structural Biology*, 118:197–219, 1997.

Index

Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation eigenständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Die Dissertation habe ich weder als Arbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht noch ist sie oder ein Teil dieser als Dissertation bei einer anderen Fakultät oder einem anderen Fachbereich eingereicht worden.

Mainz, im November 2006

Publications Appendix

I was involved in two other projects that were not presented above, but the results of the cryo-EM 3D-reconstructions can be found in the publications that are appended.

[Lieb *et al.*, 2006]

[Meissner *et al.*, 2007]