3D-Electron Microscopy of different Conformations of Cephalopod Hemocyanins

Dissertation

zur Erlangung des Grades

Doktor der Naturwissenschaften

am Fachbereich Biologie

der Johannes Gutenberg Universität Mainz

Arne Möller

geboren am 12.5.1978 in Schweinfurt

Mainz, 2009

Table of contents

Introduction

The determination of the structure of a protein is a prerequisite to understand its assembly and function, since structure and function are always coupled. The use of Transmission Electron Microscopy (TEM) of cells and macromolecular complexes bridges the gap between light microscopy and high resolution X-ray crystallography and Nuclear Magnetic Resonance (NMR) studies. The method of single particle analysis, a computer aided analysis of macromolecules with TEM, is of increasing importance, especially for flexible proteins and large complexes. Datasets with different environmental conditions can be generated and observed. By now 3D maps of large symmetric proteins can be generated routinely at a resolution below 10Å. In combination with high resolution X-ray and NMR information of related substrucutres these maps can be upgraded to yield pseudo atomic resolution.

Aerobic living organisms convert their nutrition efficiently into energy-containing ATP using O_2 as the final electron acceptor in the respiratory chain. While protozoa and small metazoa can facilitate oxygen uptake directly from the surrounding water or air by diffusion, this becomes more difficult as the mass of the animal increases. Adaptations, such as tracheas, gills, lungs and pumping within circulation systems, together with respiratory proteins, have evolved to cope with the problem of oxygen supply to the tissues. The respiratory proteins can function for oxygen storage or as transporters. They bind 0_2 reversibly, to release it locally at the consuming tissues. The most widely studied oxygen transporter is human hemoglobin, a tetramer with a heme group, consisting of a ferrous ion coordinated by four nitrogens at the center of a porphyrin ring, as active site. This ~64kDa oxygen transport protein is present at high concentration in erythrocytes. In 1960 Max Perutz solved the three dimensional structure of human hemoglobin by X-ray crystallography (Nobel Prize in 1962, together with John Kendrew (Myoglobin)).

Hemoglobins are wide spread over almost all phyla and form many different oligomeric protein assemblies, depending on the species. All of these different types have one or more heme-porphyrin active sites to which 0_2 can bind reversibly. The shape and size of hemoglobins varies considerably. In *Lumbricus terrestris* for example, the extracellular hexameric hemoglobin has a molecular mass of ~3.5 million Daltons, 12 sub-assemblies and 144 subunits, each with an oxygen-binding site (Schatz et al. 1995). But, hemoglobins are not the only group of iron-containing respiratory proteins found in animals. Hemerythrins are non-heme iron proteins found in a few marine invertebrates.

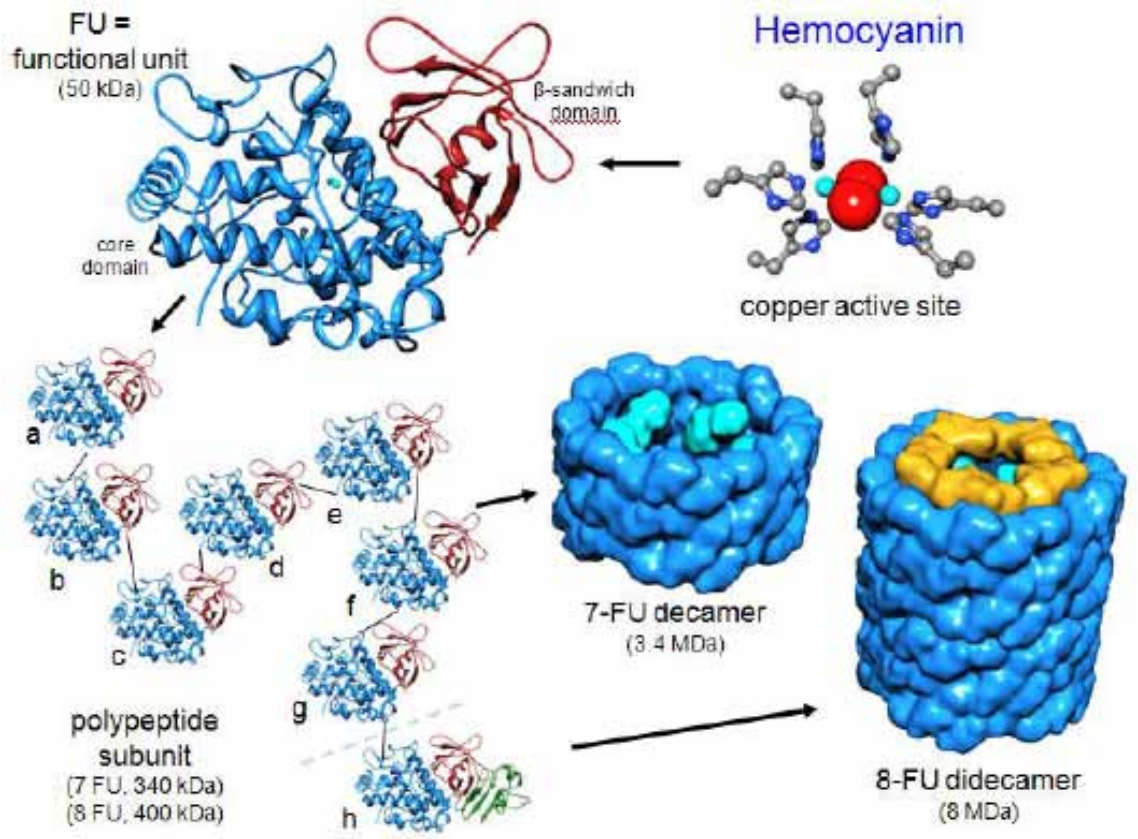


Fig. 1: Principle assembly of molluscan hemocyanins

Their oxygen-binding site is composed of two ferrous ions linked by carboxyl side chains.

The group of oxygen transport proteins dealt with throughout this study are the Hemocyanins, giant extracelluar blue copper proteins freely dissolved and circulating in the hemolymph of many molluscs and arthropods. The hemocyanin active site is composed of two copper ions coordinated by six histidines. During oxygenation the

Cu(I) state copper ion becomes Cu(II), binding the oxygen as O_2^{2-}. This complex is the reason for the deep blue color of oxygen-loaded hemolymph (van Holde et al. 1995)(see Fig.1). Due to their similarities, with respect to copper-dependent oxygen binding, the hemocyanin molecules of mollusks and arthropods were given the same name. However, they are a diphyletic group, because identity is limited to the active site organization rather than the overall protein (Markl and Decker 1992; van Holde et al. 2001). In principle arthropdian hemocyanins are assemblies of one or more hexamers, containing bean-shaped monomers of mass ~75kDa. These hexamers form multimers with up to 48 active sites (8x6mer) (Martin et al. 2007). Structurally, molluscan hemocyanins are, on the other hand, large polypeptide chains with seven or eight sequential paralogous functional units (FUs) (termed FU-a to FU-g,h), with 350-450kDa (see Fig.1). The globular FUs are connected by linker peptides (10-15 amino acids) (Lang 1988; Lang and van Holde 1991) giving the monomer a pearl necklace-like appearance. Ten of these monomers form a hollow cylindrical decamer, with a five fold rotational symmetry, immediately suggestive of underlying subunit dimerization (see Fig.1). By electron microscopy these hollow cylinders appear as ring-like molecules, with a fenestrated wall encircling an internal arc structure. Whereas in gastropod hemocyanin the c-terminal FU-h copies form a collar partially closing the decamer at one side, in *Sepia* hemocyanin the eighth FU originates from a gene duplication event of FU-d. Thus, this additional FU is in the centre of the polypeptide chain, the decamer is not closed at one side but the number of FUs contributing to the arc is doubled (Lambert et al. 1994; Oliver Lambert et al. 1994; Boisset and Mouche 2000). In gastropods and chitons two decamers can assemble face to face forming didecamers with a mass up to 9MDa and 160 oxygen-binding sites. Even higher assemblies are possible, forming long tubes, termed multidecamers (Harris et al. 1997). Because of their molecular size, stability, ease of purification and widespread distribution, hemocyanins have studied greatly by the techniques of protein chemistry and electron microscopy. In 1972 the first three dimensional (3D) model of a molluscan hemocyanin was published in *Nature* (Mellema and Klug 1972). At the beginning of my thesis the 3D-structure of FU-g of *Octopus dofleini* hemocyanin (OdH-g) (Cuff et al. 1998) as well FU-e of *Rapana thomasina* hemocyanin (RtH2-e) (Perbrandt et al. 2003) (see Fig.2) were already solved by X-ray studies. According to the blue color of the crystal of OdH-g and the colorless crystals of RtH2-e it is indicated that the structures represent two different conformations. The comparison of both models shows very high coherence. The FU is composed of two sub-domains, namely an α-helical domain (core domain), containing the FUs active site, and a six stranded β-sandwich domain. The distance between the two coppers (CuA and CuB) at the active site varies between both conformations (~3.5Å

OdH-g and ~2.8Å RtH2-e), implying that this is caused by oxygenation of the active site. Six histidines, attached to four central α-helices, coordinate the coppers. These histidines also show a compensatory movement upon oxygenation. Another general feature is a thioether bridge between a cysteine residue and the second histidine of CuA. This bridge could play an important role in cooperative modulation of oxygen affinity (Gielens et al. 1997). A strictly conserved leucine residue protrudes form the β-sandwich domain and contacts the active site at van der Waal distance. This contact might transfer allosteric signals between the two domains and modulate oxygen affinity. A difference between the structure of OdH-g and RtH2-e is a tunnel only found in the deoxy-RtH2-e.

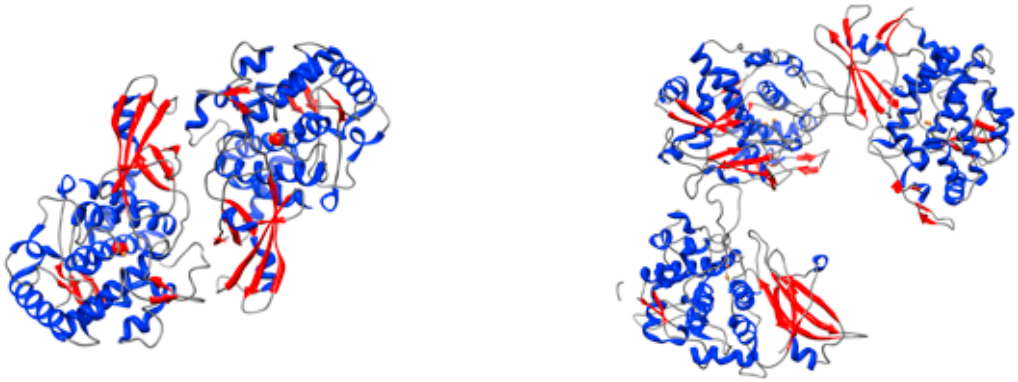


Fig. 2: Assembly of the crystallized FUs in Octopus (left) and Rapana (right)

This tunnel leads from protein surface directly to the active site of the FU and might therefore form a pathway for dioxygen molecues. Since this tunnel was absent in deoxy OdH-g it can be supposed that it is closed upon oxygenation. The arrangement of the FUs within the crystal is also different. While in OdH-g two protomers are associated as a perfect dyad symmetry in RtH2-e six protomers (two trimers) form regular cylinders (see Fig.2). Due to the high sequence identity of molluscan hemocyanins (>35%) the crystal structures can serve as templates for homology modeling of the FUs of all molluscan hemocyanins. Besides this basic building block structural investigations were made on fully assembled hemocyanins and separated dimers (Lambert et al. 1994; Oliver Lambert et al. 1994; Orlova et al. 1997a; Meissner et al. 2000; Harris et al. 2004). These led to an interpretation of the quaternary structure and the subunit organization of this molecule. However, the resolution did not allow to deduce much details with certinity.

The work presented in this thesis deals with the architecture of cephalopod hemocyanins ,and with the structural changes caused by oxygen uptake

Cryo-TEM and single particle reconstruction

The development of specimen vitrification for transmission electron microscopy (cryo-TEM) (Glaeser and Taylor 1978; Dubochet et al. 1982; Dubochet et al. 1988) was a huge step for structural biology. This preparation method allows the study of fully hydrated proteins or even cells within the electron beam (Al-Amoudi et al. 2004b). In this method the proteins are dissolved in aqueous solution and are applied to an EM-grid. This thin aqueous film is rapidly plunged into liquid ethane where it vitrifies. The fast freezing process leaves the specimen protein unharmed. An extreme cooling rate of about 100,000°/s can be reached due to the small volume and the high speed plunging (Frank 2006). This results in a freezing time, of the specimen, in the microsecond range (Al-Amoudi et al. 2004a). This is obligate if rapidly changing protein conformations, such as may occur during oxygenation, are to be trapped. However, cryo-TEM comes with some severe drawbacks. Specimen preparation is not trivial and the images suffer from low contrast. Furthermore, only proteins with a size of ~80kDa or greater can be visualized yet. Due to charging and local heating special care is also important during the electron bombardment in the microscope. An evaluated setup, using low electron dose imaging and averaging of many thousand particles can overcome this problem. In single particle analysis identical views of *single* particle images are identified and averaged. This method significantly increases the low contrast. Subsequently the relative orientation of these class-averages are estimated and projected into three dimensional space (van Heel et al. 2000), resulting in a three dimensional density representation of the molecule. With increasing computer power and better electron microscopes, near-to atomic resolution has already been reached (Hong and Zhou 2008; Ludtke et al. 2008) applying this method on symmetrical proteins. Especially for proteins that are either too large or too heterogeneous for X-ray crystallography or nuclear magnetic resonance (NMR) study, a successful future of this method is indicated. However, at the moment the biggest strength of this method is a hybrid approach. Density maps of resolutions around 10Å can easily be *upgraded* using available X-ray data and homology modeling. This procedure is very fast and straightforward. Since nearly the whole process is computer-dependent the total reconstruction procedure is likely to be automated to a significant degree within the next few years (Potter et al. 1999; Carragher et al. 2000), making it an ideal method for structure determination of proteins. Successful vitrification is a gentle preparation method that induces no artifacts or other protein transformations, as occur during chemical fixation, dehydration or staining. methods. Thus the vitrified protein

sample can be preserved in a near-to native state.

The Animals

Nautilus pompilius (Linnaeus, 1758):

Phylum: Mollusca
Class: Cephalopoda
Order: Nautilida
Family: *Nautilidae*
Genus: *Nautilus*
Species: *Nautilus pompilius*

Fig. 3: *Nautilus pompilius*

Fossils prove the existence of Nautililoidea in the Cambrian period. This is the reason why *Nautilus* is often referred to as a *living fossil*. *Nautilus pompilus* (see Fig.3) is the most frequent member of this group of mollusks. It inhabits the steep coral reefs in the South Pacific Ocean. The animals occur from Fiji in the east to the Andamans in the west and from south Japan in the north to the Great Barrier Reef in the south. At daytime they live in depth of up to 400m, but surface to 30m in the night to catch prey. Unique among all cephalopods is the helical shell. It is subdivided into chambers which the animal can use to regulate up and downward buoyancy movement by injection of fluid through a siphucle (Westermann et al. 2004). Horizontal movement is performed by a *jet-propulsion* mechanism. Water is pumped out through a muscular siphon, located under the tentacles. The animal can direct the outflow of the water, allowing *Nautilus* to swim in any direction, although the main direction is backwards.

Sepia officinalis (Linneaus, 1758), common cutllefish

Phylum: Mollusca
Class: Cephalopoda
Order: Decabranchia
Family: *Sepiidae*
Genus: *Sepia*
Species: *Sepia officinalis*

Fig. 4: A group of *Sepia officinalis*

Sepia (see Fig.4) is widely distributed around the planet. It reaches from north of the Faro islands over swathes of the eastern Atlantic Ocean, as well as the North Sea and the Mediterranean Sea. Their habitats are mainly the shallow coastal areas and natural reefs. They are mostly nocturnal and hide in the sand during the day. As a very active and intelligent predator they have the highest basic metabolic rate among all molluscs. Thus they need a very efficient energy supply.

Objectives

The subject of this study has been the hemocyanin of two different marine predators, the *living fossil Nautilus pompilius* and the cuttlefish *Sepia officinalis.* In *Sepia* the hemocyanin is composed of eight FUs. In contrast to gastropod hemocyanins the eighth FU is not appended to one end of the polypeptide chain, where it forms a "slab" closing the decamer at one side, but is included in the center of the chain. Thus, the quaternary structure is different to that of all other molluscan hemocyanins so far investigated. The hollow cylinder of *Sepia* hemocyanin is not closed, but the internal arc structure is composed of two different FUs. To understand the architecture it is crucial to know which FUs take part in this arc formation and how this influences the basic cylindrical organization. In addition to the elucidation of the subunit topology, one goal of this thesis is the construction and analysis of a *pseudo* atomic model of the hemocyanin of *Sepia officinalis.* This model could then be interpreted in terms of the subunit-subunit interface organization. Therefore, the topology of the FUs within the decamer has to be revealed. Additionally, by using the crystal structure of *Octopus* FU-g (Cuff et al. 1998) as a reference, homology modeling needs first and foremost to be done. Success would fill the missing gap and contribute to the full understanding of the molecular architecture of molluscan hemocyanins in general.

During oxygen-binding hemocyanins exhibit strong allostery and cooperativity. While this cooperativity is not as strong in *Nautilus* as in some arthropodan hemocyanins, some structural changes have been claimed to occur. In the present study the structural changes of a molluscan hemocyanin has been observed by cryo-TEM. Therefore two independent single particle reconstructions of hemocyanins, prepared under oxygenating and deoxygenating conditions, were prepared. At a subnanometer resolution the molecular model (Gatsogiannis et al. 2007) can be applied and information of how allosteric conformational changes occur could be gathered. Such a pseudo atomic model allows to directly investigate inter FU-contact organization and the comparison between the models of both states will indicate the differences. Any cooperative signaling of the molecule has to be transferred between neighboring FUs *via* their interfaces. Thus a careful investigation of residues taking part in interface organization is obligate. Together with correlating densities, deduced from the single particle analysis, which bridge the gap between the two FUs, these residues might indicate putative mechano-chemical interaction occurring during oxygenation. Of special interest are charged and hydrophobic residues because these can interact via salt bridges, binding of certain ions and hydrophobic attraction. This structure to function approach could give insight into the cooperative behavior of a molluscan hemocyanin.

Materials and Methods

Since the problem of this work is mainly biological I will only briefly describe the methods of Transmission Electron Microscopy (TEM), single particle analysis and interpretation, including homology modeling and rigid body fitting in this chapter. Detailed information can be found in corresponding literature (van Heel et al. 2000; Frank 2006).

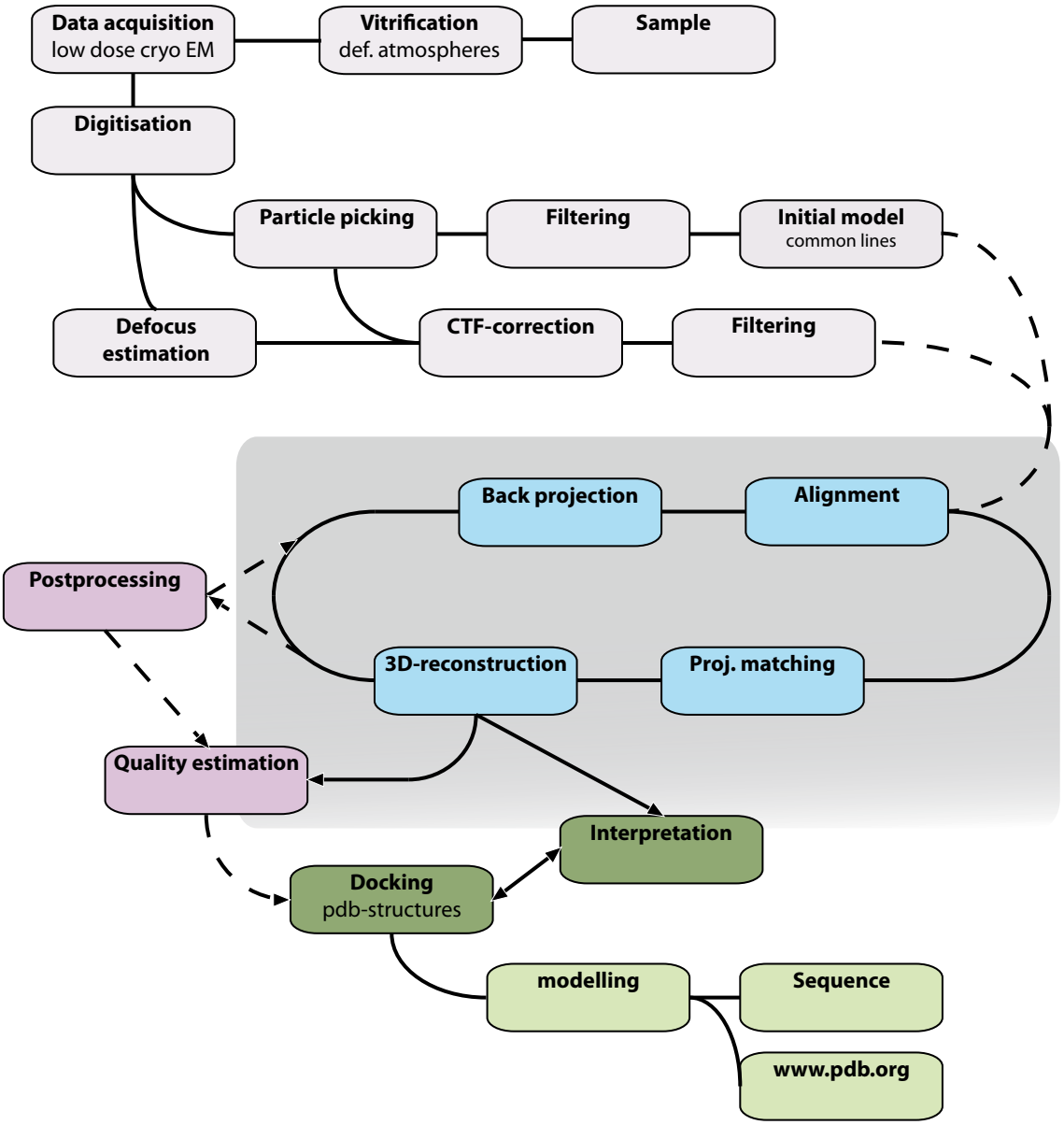


Fig. 5: Flow chart and approach of the processes involved in the single particle 3D-reconstruction. The iterative refinement procedure (grey field) was repeated many times.

Preparation of the sample

The Animals

Three living *Nautilus pompilus* were bought at a local pet shop (ZOOPALAST Wiesbaden). The hemolymph of *Sepia officinalis* was provided by Dr. Felix Mark at the Alfred-Wegener-Insititut (AWI) Bremerhaven.

Nautilus was kept in a seawater aquarium at 15°C with alternating day (faint light) and night (complete darkness) 12h periods. The animals were fed on shrimps and cod, a calcium carbonate solution was added frequently to support shell formation.

Purification of the hemocyanin

For the extraction of the *Nautilus* hemolymph a non-lethal method was applied. The living animal was cooled on ice for 15 minutes to immobilize and anesthetize it. A hole, 0.5 mm in diameter, was drilled through the shell at the area of the gills, using a PROXXON MICROMOT 50/E© electric drill tool. After that an injection needle, pointing into the gills of the animal, was then used to withdraw about 1ml of hemolymph. It is likely that this method was applied for the very first time in our laboratory and successfully performed on the same animal several times. To achieve sufficient material for the whole project, the hemolymph of two different *Nautili* was used. The hemocylymph of *Sepia* was kindly provided by Dr. Felix Mark at AWI Bremerhaven.

Hemocyanin purification was done in the same way for both animals. Directly after the extraction of the hemolymph a protease inhibitor (Pefabloc® SC-Protease-Inhibitor) was added to the sample and any cellular components were removed by centrifugation (Sorvall RC-5B, Du Pont Instruments, Frankfurt) at 1000 x g for 5min. After this the same protease inhibitor was added again. Overnight, the hemocyanins were sedimented using an ultracentrifuge (Airfuge Ultracentrifuge, Beckman Optima L-70, Munich) at 30,000 x g. The blue protein pellet was then resuspended with Tris-saline buffer (0.05 M Tris, 0.15 M NaCl, 5 mM MgCl2, 5 mM CaCl2) but at slightly different pH-values (NpH-oxy pH 7.4, NpH-deoxy pH 6.4, SoH pH 8.2), taking into account the specific oxy/deoxygenation requirements.

Vitrification

Prior to vitrification the hemocyanin concentration was measured using absorption spectra. The measurements were done in the range of 220nm to 600nm using an Ultraspec 3100 Pro UV/Visible Spectrophotometer. A typical hemocyanin delivers two characteristic absorption maxima, at 280 and 340nm. While the first is typical for all proteins the second is due to the oxygenated copper active site of the hemocyanin (Van Holde 1967). Thus this method can be used to estimate the concentration of the resuspended hemocyanin.

All vitrification preparations took place within one week after hemocyanin purification, to ensure a very fresh specimen. The proteins were always kept at 4°C and never frozen prior to fixation. The samples were plunged into liquid ethane using a locally designed and constructed plunge freezer (see below). It was essential to conserve the specimen in a defined atmosphere 24h prior to plunge the preparation. Therefore it was put in a fridge and a gentle, humidified, continuous gas flow of either 25%O_2/75%N_2 NpH-oxy or only N_2 NpH-deoxy was blown over a 10-16µl sample in a small reaction container. Because the SoH-preparation also served for another experiment, the atmospheric composition was 4% O_2 and 96% N_2. This overnight pre-incubation ensured that the solution was in equilibrium with the desired oxygenation state. The specimen was then directly transferred inside the plunge-freeze robot, the atmosphere chamber of which also contained the desired gas mixture. 2 µl of a 2mg/ml concentration of the protein was applied on each Quantifoil grid (Quantifoil Micro Tools, GmbH) and 2-sided, 2s blotting was applied. The frozen grids were directly transferred under liquid nitrogen to small grid boxes and stored in liquid nitrogen in a Dewar flask.

This plunger–freezer (see Fig.6) was designed during this work by a co-worker in the Institute of Zoology, University of Mainz, Dr. Frank Depoix, and was an essential prerequisite for my doctoral studies. This semi-automated plunge freezer is equipped with a tube-like atmospheric chamber surrounding the main plunging device. The atmospheric composition is controlled by mixing gases from high pressure gas cylinders (premixed gas can alternatively be used). The oxygen concentration, as well as the humidity, is measured by probes inside the chamber. Since the humidity of cylinder gas is below 4%, the airflow is directed through a water bottle. This humidifies the gas up to 80%. If the water bottle is missing the sample droplet evaporates so quickly that no preparation is achievable. Two seconds of blotting was applied using the automated

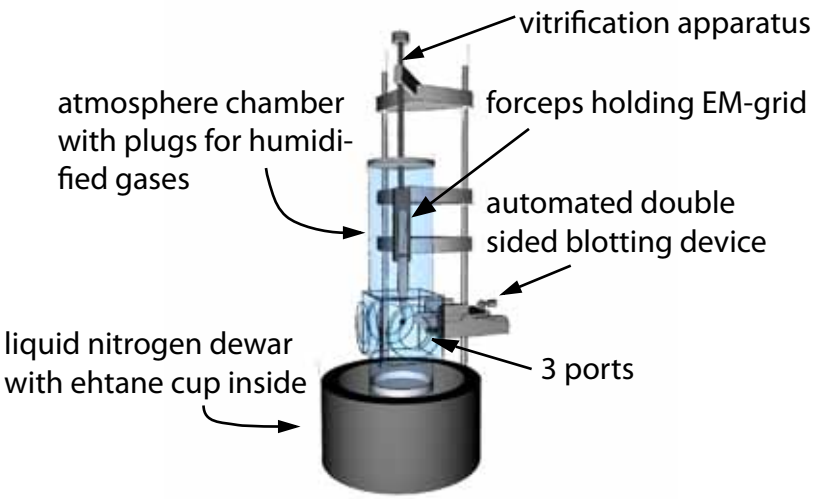


Fig. 6: Sketch of the homebuilt plunge freezer with atmosphere chamber. The atmospheric composition is controlled by mixing gases from high pressure gas cylinders (premixed gas can alternatively be used). Probes inside the chamber can measure humidity and oxygen concentration. An automated blotting device allows time controlled two sided blotting of the specimen. Three ports allow easy acess to the chamber. In case of atmosphere specific preparation the specimen will be docked at one of these ports.

double-sided blotting system. Detailed information on this plunge freeze robot will be published elsewhere.

Data acquisition

Tecnai F30/20

The cryoTEM images were taken on a Tecnai F30 (*Nautilus*) (see Fig.7) (Dr. Frank Depoix) and a Tecnai F20 (*Sepia*) (by myself); both microscopes where operated at 200kV and a magnification of 49.000. The datasets where recorded on Kodak SO 163 film but in contrast to NpH the SoH dataset was recorded using low back-scattering holders, designed by the FEI-Company. These holders have a large rectangular hole in the center, reducing back-scattering effects from the holder material. The overall effect produces lighter, less noisy images. The defocus range used was between 2.8 and 0.8µm .

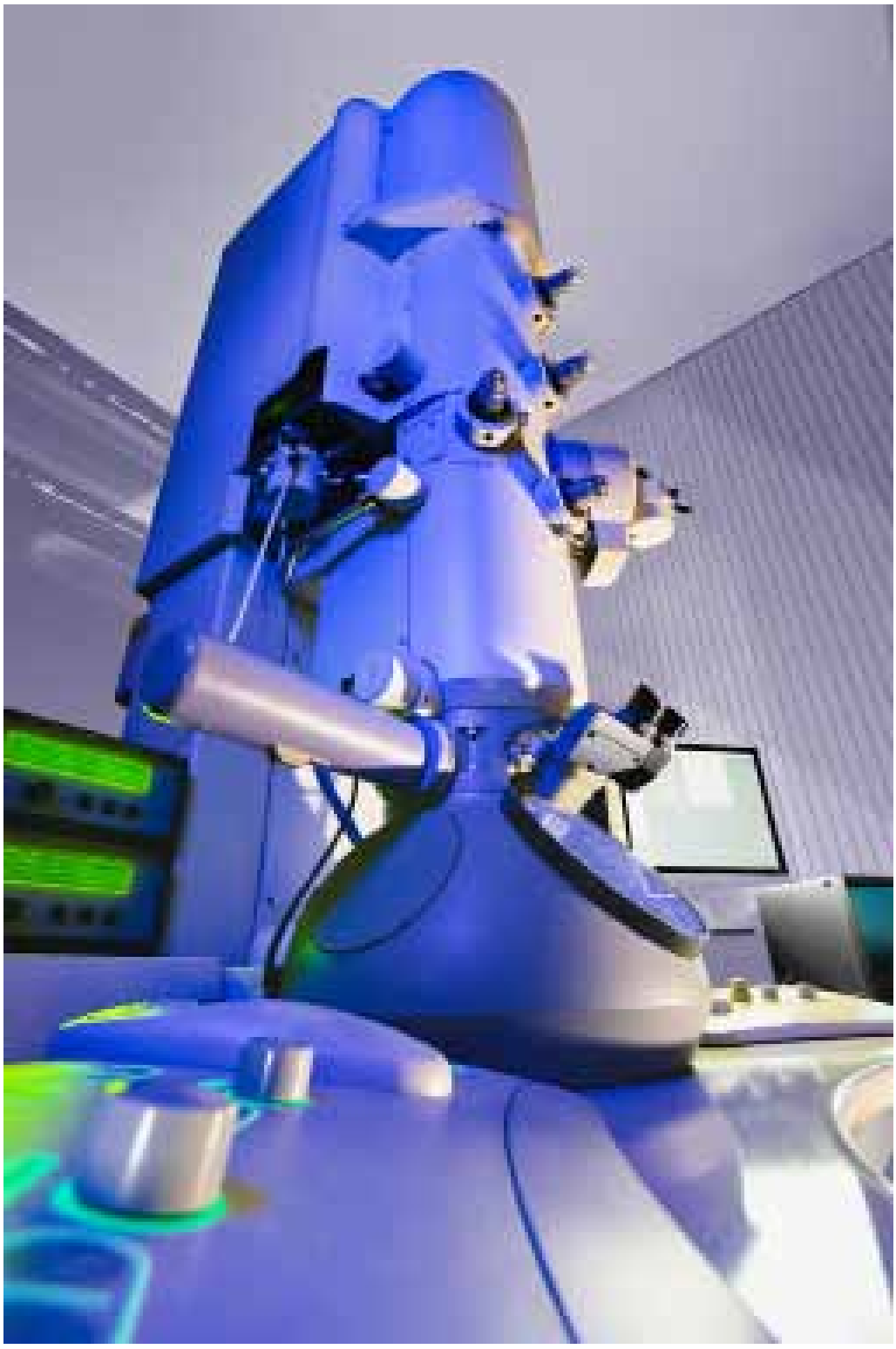


Fig. 7: Tecnai 30 FEG TEM at the Johannes Gutenberg University Mainz (Picture by Christoph Kuehne)

Scanning and digitization of micrographs

The cryoTEM negatives were scanned using a HEIDELBERGER drum scanner (Primescan) and saved as 8 bit grey scale Tiff-files, using the Linocolor software. The *Nautilus* dataset was digitized with 1.05Å/pixel (5.14µm step size) and *Sepia* with 1 Å/pixel (5µm step size).

Computer-Hardware

The image processing was performed on a "Beowulf" cluster equipped with 20 blades (transtec, Tübingen, Germany). Each blade had two AMD Opteron dual core processors (2.4GHz) and 16GB RAM. (see Fig.8)

Fig. 8: Transtec "Beowulf" cluster at the institue of Zoology Dep.II Johannes Gutenberg University Mainz (picture by Christoph Kuehne) →

Image Processing

Contrast and Contrast Transfer Function

Contrast in an electron microscope results from specific electron-specimen interactions. Generally these are due to elastic and inelastic scattering. While the elastic scattered electrons interact with the specimen without energy transfer the inelastic scattered electrons loose energy and are scattered by small angles. The energy is then transferred

to X-ray, secondary electrons or energy such as light. All effects of inelastic scattering contribute to background noise and are in principle unwanted. Furthermore the resulting energy causes problems in cryoTEM, such as local drift and ice melting. The elastic scattered electrons are spread at wide angles and can therefore yield high resolution information (see Fig beam sample interaction). While negative stained samples alter the amplitude of the electron beam passing through it (amplitude objects), in ice-embedded specimens only the phases are shifted (phase objects). This phase-variation is not detectable because of the high frequency of an electron beam accelerated by 200kV. However, applying an aberration, ie bringing the projector lens out of focus, can reveal this phase information. This is because the phase in a defocused electron microscope is located in the back focal plane. Thus, the phase information in the defocused electron microscope image is no longer zero everywhere, making the phase information visible. A delay of λ/4 (90°), with respect to the phase at the optical axis, leads to ideal (positive) phase contrast in the back focal plane of that particular spatial frequency. If an underfocus is combined with the spherical aberration of the microscope, this leads to phase shift of exactly λ/4 over a large range of spatial frequencies. Essentially, this combination makes the electron microscope perform as a phase contrast microscope.

The actual electron image is convoluted with the Point Spread Function (PSF) of the microscope. The PSF can be defined by acceleration voltage, defocus, electron wave length and the spherical aberration coefficient. The Fourier transform of this PSF is the contrast transfer function (CTF) which is specific for each electron microscope. The CTF oscillates between +1 and -1, causing more and more phase reversals with increase of the higher spatial frequencies. This leads to a misinterpretation of high resolution data, but can be corrected if the exact zero crossings of the CTF are known. Effects like beam incoherence, drift and the finite electron source size dampen the CTF to zero at a certain point. This limits the resolution of the microscope.

Particle selection

All particles were selected manually using the EMAN tool *boxer* (Ludtke et al. 1999). To enhance visibility of the particles the particle images were binned by a factor of four and low-pass filtered using the IMAGIC-5 tools coarse-image and filter-image. Although many different programs now provide automated particle selection, I decided against it for two reasons. In the case of SoH and NpH tilted views were rarely picked by the software (EMAN, Signature), but are very useful for the reconstruction. The second reason is that, on average 70 particles were picked from a single electron micrograph.

Thus, a quick manual “pick” was simply faster then an automated ”pick”. The coordinates of the selected particles were multiplied by four (the coarsing factor) and extracted into boxes of 512x512 pixels using the IMAGIC-5 tool subfile-images. After CTF-correction the boxes were subfiled to 400x400 pixels to increase computation speed.

Defocus estimation

As described above in an underfocus is applied to visualize the particles. It is crucial to exactly know the values of this defocus in order to correct the data for it. The exact estimation of the defocus values was performed using the ctffind3 package (Mindell and Grigorieff 2003).

Correction of the defocus

In this work the CTF was corrected by flipping the phases at the negative lobes, only. The phase flipping procedure was performed with IMAGIC-5 (van Heel et al. 2000).

Single particle reconstruction

Initial model and the iterative refinement

The initial model was acquired using the standard procedures implemented in IMAGIC-5. This means an initial search for relative orientation of three different views using sinogram correlation implemented in the IMAGIC-5 tool *euler-angular-reconstitution*. The bandpassfilter (see below) applied to the single particles differed significantly from the one used for the iterative refinements, because for this step no high resolution data is needed.

Preprocessing

The raw data particles were filtered and normalized prior to the analysis. Therefore a bandpassfilter was applied, a combination of a wide Gaussian low-pass filter and a narrow high-pass filter. The suppression of low frequencies increased the accuracy of the alignment and avoided misalignment to background features as well as dominance of

low resolution information. The high frequencies contain the high resolution information of the dataset, but also much noise. Since the purpose of an initial molecular model is only to produce a starting position, these high frequencies were also suppressed significantly, to a value of 30Å. Normalization was done using a new σ of 10. Finally a circular, soft mask was applied to the single particles. The filtering and normalization of the raw particles was done with the IMAGIC-5 command *incore-prepare-filtered*. After the initial model was generated all images were CTF-corrected and filtered again. With increasing resolution the borders of the filter were slowly shifted towards higher spatial frequencies in a step-wise procedure, until finally high frequencies at a value of 6Å were allowed.

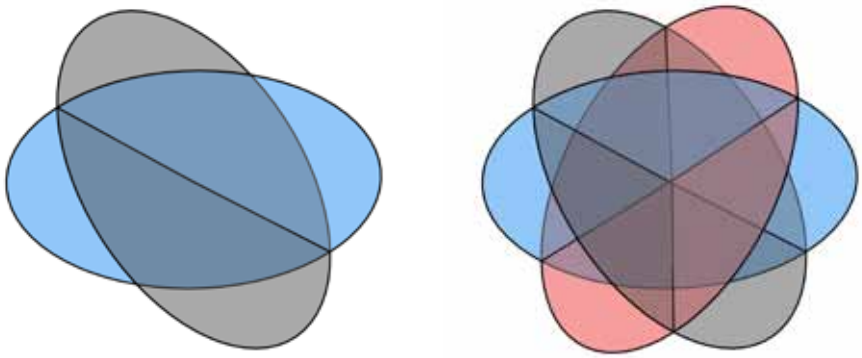


Fig. 9: Common lines of round discs. The combination of two discs (blue and grey -left side) is not sufficient. The third disc (red- right side) fixes the system.

Alignment and Classification

Prior to the Multivariate Statistical Analysis all particles were centered to the rotational sum of all images using the IMAGIC-5 command *center-images.* As a model had already been generated, it was be used, instead of the rotational sum, as reference for a *multi-reference-alignment* (m-r-a). In this procedure all single particles were correlated with 2D reprojections of a 3D-object. Since we found many different orientations in a dataset many references were needed also. The alignment

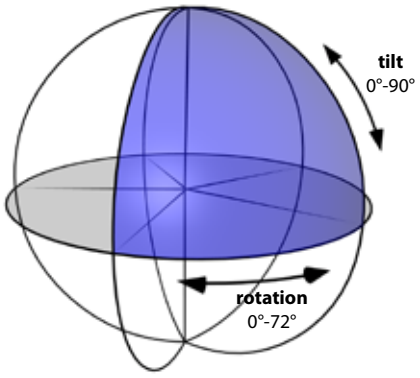


Fig. 10: Asymmetric triangle (blue) of a D5 symmetric object. In case of a C5 symmetry the blue triangle is expanded over 180° tilt angles

correlated all images with all references. The result was a correlation coefficient which was used to identify the “best fit”. The images were then aligned (shifted, rotated) with the reference to which they share the highest correlation coefficient. The alignment was the most time consuming procedure in the iterative process. Therefore I wrote a simple distribution script that increased the speed of this step dramatically (see below).

Multivariate statistical analysis

Multivariate-Statistical-Analysis (MSA) (also referred to as Multivariate data analysis (Frank 2006)) is an integral procedure in single particle analysis. It’s basic purpose is to classify particles depending on similarities within the densities of the distributed pixels. Similarly distributed grey values within an image should also show the same view of the same object and should therefore be incorporated within the same class. Another useful aspect of MSA is the data-compression that comes along with this method. Using hierarchical classification, classes of images with the same orientation can be generated. This classification procedure optimizes the relationship between intra class similarity and inter class difference (Ward 1982). The summation or averaging of these images leads to an improvement in signal to noise ratio (SNR) because the discrete information within the particles predominates, while the random noise is averaged out. (*see* (Van Heel and Frank 1981) and (van Heel et al. 1996) for further details)MSA was performed using the IMAGIC-5 commands *msa-run, msa-classify-images* and *msa-sum-classes.* In case of the initial model relatively few eigenimages were taken into account (3-8), to concentrate on major differences. To enhance the signal, between 50 and 500 single particles formed one class.

Common lines

To obtain a 3D reconstruction, at least three different views of a particle are needed. Since ice-embedded proteins can be oriented in a random manner the user has to determine their relative spatial orientations, ie their Euler angles, in order to obtain a correct reconstruction. (see Fig.9)

The Euler angle determination of an image, in this case a classsum, is based on the angular reconstitution technique (Van Heel 1987). According to the common-line theorem, each 2D projection of a 3D object has a 1D line projection in common (DeRosier and Klug 1968). The angle between such common lines reflects the relative Euler-angle orientation of the projections to each other. IMAGIC-5 uses sinogram (line integral) correlation procedures

to find these Euler angles. Therefore a pair of sinograms of 2D projections is compared line by line in a correlation function. At the point of highest correlation is a common line. In case of multiple symmetries, more than one peak can be observed.

Symmetry

The symmetry of an object can be exploited in a way that it artificially increases the number of particles. Thus, one single image, or class of images, can be used more than once for the reconstruction process, as long as the symmetry results in an increase of signal. NpH exhibits a five fold rotational symmetry (C5), whereas SoH has an additional symmetry axis, orthogonal to the central symmetry axis (D5) giving the structure a 10-fold symmetry. (see Fig.10)

Projection matching

The projection matching (Penzcek et al. 1994) process is an alternative to the orientation definition using common lines. The idea of it is simple. Each image is assigned the same Euler angles as the best matching reference. All images with same angles are then summed within the same class. This procedure was used throughout all iterative cycles, after the initial model had been created.

3D-reconstruction

Molecular 3D reconstruction follows the central projection theorem. It states that the projection of a 2D function can be obtained as the 1D Fourier Transform (FT) of a central section through its 2D FT (Radon 1917; Bracewell and Riddle 1967; DeRosier and Klug 1968). Therefore, accurate back projection of class-averages lead to the 3D structure. This process is performed in IMAGIC-5, by applying the “exact filter back projection” (*true-threed-reconstruction*) module (Harauz and Van Heel 1986; Radermacher 1988; Radermacher et al. 1994; Schatz et al. 1995). This filtering helps to avoid the over weighting of low resolution information, thereby enhancing high resolution information. Overall NpH-oxy 6729, NpH-deoxy 9129 and SoH 8862 superior single particles were used for the reconstructions.

The alignment using subsets

Within a typical IMAGIC-5 m-r-a, each image is correlated to each reference, to identify the best image/reference correlation. A typical final refinement iteration of the datasets presented here consisted of ~10,000 particles and ~8,000 references. This means 80,000,000 comparisons are performed within only one iteration. On the 80core Beowulf cluster (see above) this process took days, if not weeks, depending on the box size and Å/pixel size. However if the alignment is already stable, most of these comparisons are dispensable since the particles are already well-oriented. To increase the speed of the alignment procedure I wrote a script which splits the set of references into subgroups, depending on their Euler angles. The dataset was then sorted into groups of references, within which the single particle showed the highest correlation in the previous alignment. The numbers of subsets were altered between two iterations, to allow overlapping at the edge of a reference subset. In practice, a division of the dataset into groups of 100, or even more, was practicable. This in turn led to an increase of processing speed by nearly a factor 100. In fact, the alignment procedure was reduced from a couple of days, to a couple of hours. Significantly, this treatment enhanced the overall work-flow of the whole EM-group.

Quality estimation

The quality estimation process include the Fourier shell correlation (FSC) method (Harauz and Van Heel 1986), where two 3D-reconstructions created from halves of the dataset are correlated. The resolution given in this work is based on the 0.5 criterion (Bottcher et al. 1997), which is considered rather conservative. For reasons of better comparison, the ½ bit criterion is although displayed within the FSC-graphs (van Heel and Schatz 2005). Additionally the maps where compared with artificial models generated from rigid body-fitted pdb-models, by eye. Within the present work, special emphasis was placed on the appearance of secondary structure elements.

Postprocessing

In this process, a wide soft mask was applied onto the 3D model to exclude noise. Furthermore, the model was filtered to the estimated resolution. A lowcut filter (sharpening) was applied to enhance high frequency information. This high resolution information is sometimes missing in the surface representation, due to over-representation of low resolution information. This was either done using the IMAGIC-5

command *threed-filter-image* or the EM-BFACTOR software (Fernandez et al. 2008).

Homology modelling

In case of NpH the homology models already existed and were not further modified (Gatsogiannis et al. 2007). In case of SoH both isoforms were modelled strictly following the procedure used for NpH. The software used for this purpose was modeler 9v1 (Fiser and Sali 2003), with the crystal structure of *Octopus* FU-g as reference (Cuff et al. 1998).

Interpretation of the structures

The interpretation of the structures founded on rigid body fittings of modelled pdb structures. Automated fitting was performed with Chimera. The interpretation of the data was performed in close correlation with the densities gained from single particle analysis and the fitted pdbs. Special emphasis was directed to the interfaces between the FUs.

Software

Adobe Creative Suite (v. 3+4) is a professional software package and includes Adobe Photoshop, Adobe Illustrator and Adobe InDesign. During the course of the work these programms were used to create images, design posters and presentations and to edit the thesis. (http://www.adobe.com)

Amira-**Resolve RT** (v. 3 to 4.1) is a professinal 3D-viewer with many applications, such as tomography and segmenting. It was used to visualize the 3D-maps. (http://www.tgs.com)

Cinema 4D (v. 10) is a professional animation and 3D-graphics software package. It was used to render images of the 3D models and to create animations for presentations. The software allows to render *Virtual Reality Objects* a kind of movie in which the user can interactivley rotate the 3D object. (http://www.maxon.net)

ClustalW and **ClustalX** are free software packages that carry out automatic alignments of sequences from DNA or protein. These programms where used during the homology modelling processes. (http://www.clustal.org)

CTFFIND3 is a free software package which can determine the defocus and astigmatism of an electron micrograph. This package was used for all defocus estimations performed during this work. (http://emlab.rose2.brandeis.edu/grigorieff/download_ctf.html)

EMAN (v. 1.7) is a free software package capable to carry out a complete single particle analysis. During this work the particle picking tool *boxer* and the pdb-data to density conversion tool *pdb2mrc* was frequently used. (http://blake.bcm.tmc.edu/eman/)

Genedoc is a free software package used for visualization and editing of protein and nucleic acid sequences. It was used during the homology modelling process. (http://www.nrbsc.org/gfx/genedoc)

IMAGIC-5 was the main software package used during this work. It is a modular single particle analysis program which was used for all key steps in 3d reconstruction. (http://www.imagescience.de)

Linocolor (v. 6.0.12) is a software package desigend to operate the PRIMESCAN drumscanner. It was used to scan all micrographs used in this work

Modeller (v. 9v1) is free and was used to perform the homology modelling process. It performs the comparative modelling of the target structure (protein sequence) and a related reference (pdb). (http://salilab.org/modeller)

Molprobity is a free assescible web server which was used to estimate the quality of the homology models. (http://molprobity.biochem.duke.edu)

NMFF is a free software package that can carry out flexible fittings of pdb-structures into 3d-density maps. (http://mmtsb.org/software/nmff.html)

SITUS is a free software package for automated ridgid body fitting of pdb-structures into density maps. (http://situs.biomachina.org/)

SWISS-MODEL is a free and automated homology modelling web server. (http://swissmodel.expasy.org/SWISS-MODEL.html)

UCSF Chimera is freely available and was used throughout the whole work. It is an interactive molecular graphics system with many special applications (such as rigid

body fitting, segmentation of density maps, etc...) . (http://www.cgl.ucsf.edu/chimera/download.html)

The "Beowulf" cluster is operated on a Debian 3.1 system, the desktops PCs run with SuSE Linux (v. 9-10.3), Windows XP and VIsta and MAC OSx.

Results and Discussion

Structure of *Nautilus pompilius hemocyanin* (NpH)

In 2006 our group published two papers concerning the structure of NpH (Gatsogiannis et al. 2007; Meissner et al. 2007), to which I contributed with the results of my Diploma thesis (Arne Moeller 2005). The first presented a comparison between NpH and a didecameric hemocyanin from the European marine abalaone *Haliotis tuberculata* (HtH). While the HtH molecular volume resulted from a refinement of the data published in 2000 (Meissner et al. 2000), the NpH reconstruction was derived from the initial dataset prepared from specimens in a specified oxygenation atmosphere for a molluscan hemocyanin (see Fig.11). Due to imprecision during the specimen preparation these oxy and deoxy datasets were problematic, and not useful for image comparisons. In order to gain a high resolution map the two sets were combined. This combined stack served for the reconstruction that led to the molecular model published in 2007 (Gatsogiannis et al. 2007), but this reconstruction has no specified state of oxygenation.

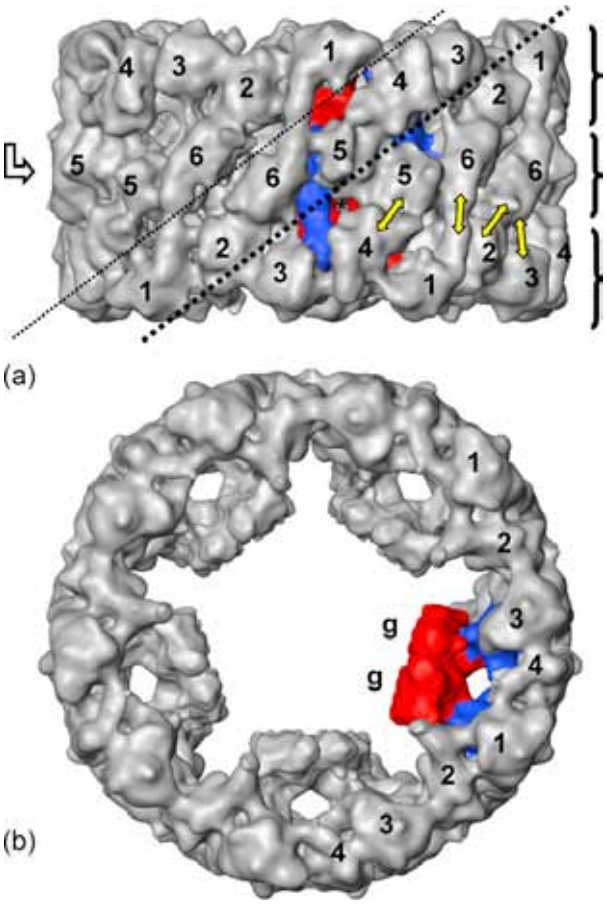


Fig. 11: Strucutre of NpH reconstruction with wrongly applied D5-symmetry (a) side view, (b) top view. Note in (a) the major groove (strong diagonal line) with two large windows, the minorgroove (weak diagonal line) with several narrow windows, and the three wall tiers (brackets). Yellow double arrows, bridges between peripheral and central tiers. Wrinkled white arrow, plane of dyadsymmetry (taken from Meissner et al. 2007).

In the first publication the “anchor” feature (see Fig.11+12) within the hemocyanin was

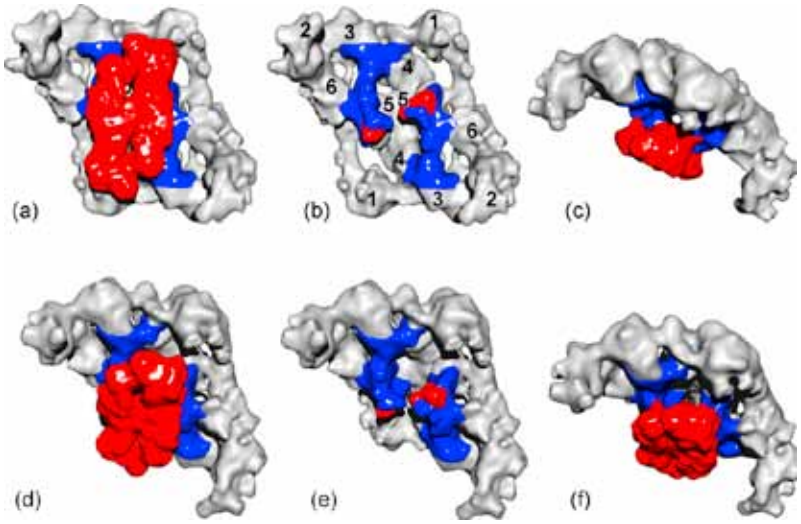


Fig. 12: The anchor structure and the collar with D5-symmetry in a wall segment. (a) Wall segment viewed from inside the cylinder (red: arc; blue: anchor). (b) The same view as in (a), with the arc removed; note that each anchor bifurcates and has two contacts. (d) The same fragment in an oblique view. (e) The same view as in (d), with the arc removed. (f) A slightly different oblique view to reveal the two bridges between arc and wall (taken from Meissner et al. 2007).

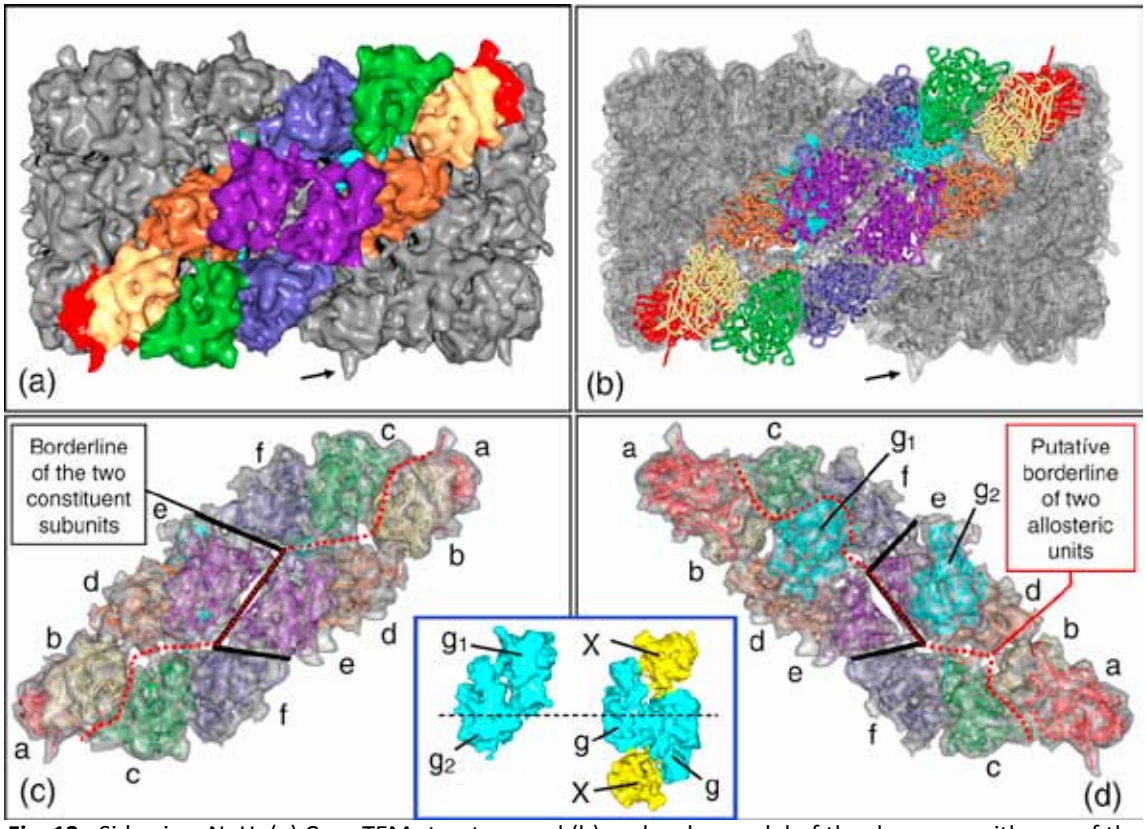


Fig. 13: Side view NpH. (a) Cryo-TEM structure and (b) molecular model of the decamer, with one of the five subunit dimers highlighted. The following colour code is used throughout: FU-a, red; FU-b, yellow; FU-c, green; FU-d, orange; FU-e, purple; FU-f, blue; FU-g, cyan. (c) and (d) Cryo-TEM structure and molecular model of the extracted subunit dimer, viewed from (c) outside and (d) inside the cylinder. The FU topology, the pathway of the constituent subunits and the putative allosteric units are also indicated. The inset shows the arc architecture of the structure as obtained by imposing C5 symmetry (left), and the doubled arc of the wrongly applied D5 symmetry (right); both structures are viewed from inside the cylinder (taken from Gatsogiannis et al. 2007).

described, as a connection between FUs a, b, e and d, for the first time. At this time it was wrongly thought to bridge the FUs a and b (since the topology was not identified, but named FU-3 and 4 in this paper) with the collar complex. This misinterpretation was caused by the finding that the molecule has C5 symmetry and not a D5 (a symmetry used in all earlier reconstructions of cephalopod hemocyanins). If the correct symmetry is applied the anchor can be identified as the connection between the four FUs a, b, d and e. This structure is well described in the second publication and also appears in the KLH1 molecule (Gatsogiannis and Markl 2008) (see Fig.13).

The problematic feature resulting from D5 symmetry, which is in my opinion the reason for its being used in all prior reconstructions, is that the wall is indeed D5 symmetrical, while the arc is not. During the iteration process the symmetry of the wall, having a strong signal in EM single particles, is over-emphasized and the resulting molecular model then exhibits a D5 symmetry. Even more important is the fact that once the D5 symmetry was applied there is no way to correct this failure using the commonly available reconstruction and alignment techniques for single particle analysis. This is due

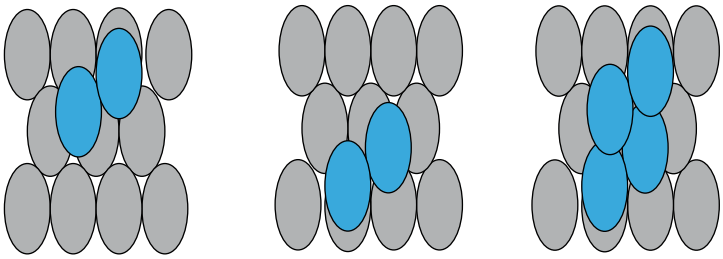


Fig. 14: Schematic view of the problematic arc architecture. From left to right: C5 symmetrical wall segment with collar (cyan), same object rotated by 180° and superposition of both, now exposing a D5 symmetry (also Compare with Fig.13) .

to the fact that the C5 symmetric map has two identical correlation values with the D5 map. These are the correct orientation of a side or tilted view and the same view turned by 180° around the dihedral axis (see Fig.14).

One way to correct this wrong assumption is to create an artificial C5 model and use this as reference for further iteration. This is the way we performed all further NpH reconstructions, which has led to the more accurate final molecular models.

Structure and molecular model of *Sepia officinalis* hemocyanin (SoH)

The single molecule of molluscan hemocyanins has either seven or eight FUs. These FUs are arranged in a linear pearl necklace-like appearance along the polypeptide chain, from the C- to N-terminus. Each FU has a molecular mass of ~50kDa and is connected to its neighboring FU by a short linker peptide. Though SoH has eight FUs, it's polypeptide chain composition is unique among all other molluscan hemocyanins. In contrast to the other eight FU-type hemocyanins, the eighth FU SoH is in the middle of the polypeptide chain, as a result of a gene duplication event of FU-d, and is not located at the end. The eight FUs are termed: N'-a-b-c-d-d´-e-f-g-C'. Prior to this work, it was unclear how FU-d' was integrated into the cylindrically shaped decamer. Thus, the main goal was to reach a sub-nanometer resolution with a molecular reconstruction in order to be able to perform high quality rigid body fittings of homology modelled crystal FU structures, to reveal the true topology. This approach has already been successfully applied to other hemocyanins, thus solving their FU topology and providing a *pseudo* atomic model (Gatsogiannis et al. 2007; Martin et al. 2007; Gatsogiannis and Markl 2008; Markl et al. 2009).

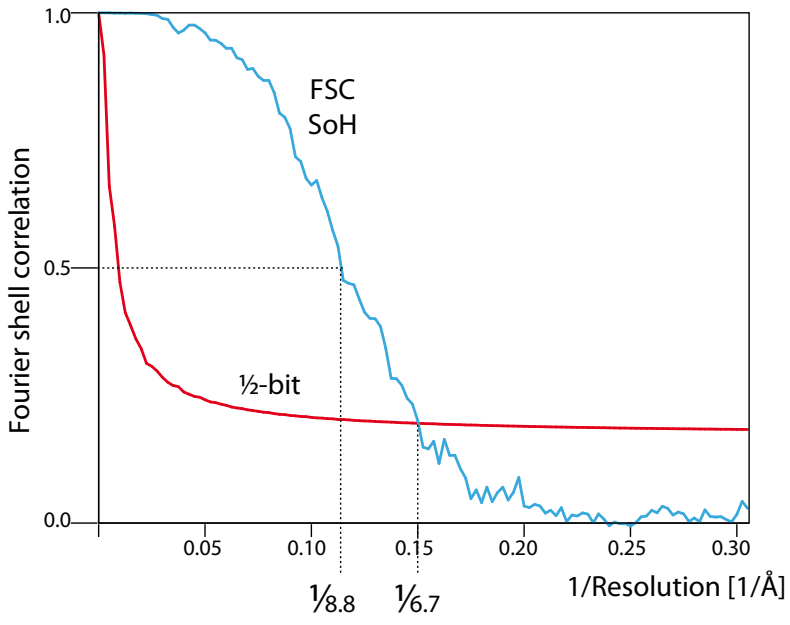


Fig. 15: The FSC-curve of the SoH reconstruction. In this work the rather conservative, but commonly used, FSC=0.5 criterion was denoted. For comparison with the publications of NpH and KLH1 (Gatsogiannis et al. 2007 and Gatsogiannis & Markl 2008) the ½-bit values are also shown. (8.8Å (FSC=0.5), 6.7Å (½-bit))

Using 8862 randomly orientated, superior selected particles from 168 cryo-TEM micrographs, a stable volume at 8.8Å resolution (FSC=0.5) was obtained (see Fig.15+16). Compared to the reconstruction of SoH published previously (Boisset and Mouche 2000), the present reconstruction shows significantly greater detail. All FUs are

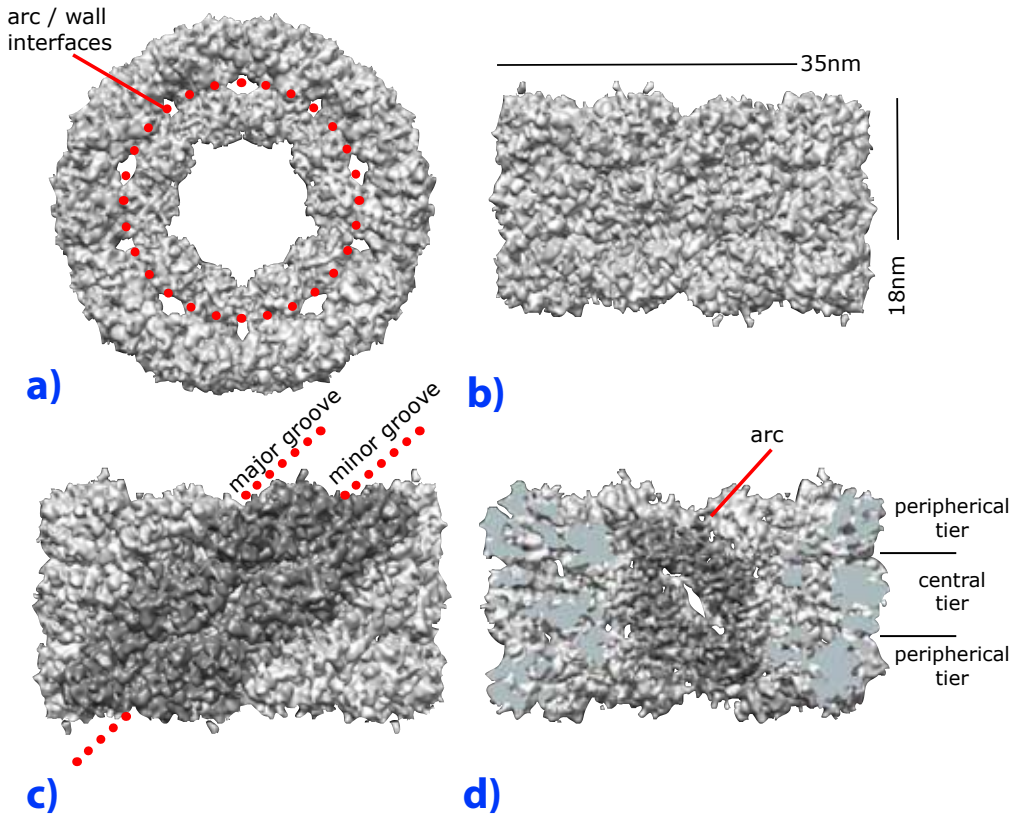


Fig. 16: 8.8Å surface representation of SoH. Top view (a), side view (b), side view with grey dimer (c) and cut open side view with colored arc structure (d).

clearly discernable and secondary structure elements are now visible. Especially, the six-stranded beta sheet domain as well as the exposed helix α15 is clearly cognizable (see Fig.18). This was already the case in the other two molluscan hemocyanins, also published at a sub-nanometer resolution (Gatsogiannis et al. 2007; Gatsogiannis and Markl 2008). The hemocyanin of *Sepia* is an oligomer of five molecular dimers, arranged as a *quasi* helical cylindrical structure, exhibiting D5 symmetry. Each SoH dimer is composed of two polypeptide chains with 8 globular FUs, each possessing an oxygen-binding active site each. The protein monomers are oriented in a parallelogram-like style (Meissner et al. 2000). Two of the FUs contribute to the inner arc structure, while the other six, are used to form the wall of the SoH cylinder. In SoH, sequence information is available for the two known isoforms (ACC. DQ388569). However, the sequence identity is very high. Since the isoforms were not individually purified for the vitrified specimen used for this study, it cannot be stated with any certainty whether the molecular volume presented here accords to a mixture of both isoforms, or whether it tends towards one.

In previous studies the principal architecture of molluscan hemocyanins has already

been shown in detail (Gatsogiannis et al. 2007; Meissner et al. 2007; Gatsogiannis and Markl 2008). The six first FUs of each polypeptide chain (abcdef) built the “wall” of a hollow cylinder, while the seventh (FU-g) forms an internal structure (termed the “arc”). If present, the eighth FU (h) forms a “slab” partly closing the decamer at one side. While the “wall” exhibits a D5 symmetry the complete cylinder has C5 symmetric, due to the off-set location of the “arcs” and “slabs”. The analysis of the inter-FU interfaces, together with the conserved molecular identity of the “wall”, revealed that the architecture is always strikingly similar, even though the total number of FUs varies. This is also true for SoH. A comparison of the SoH wall with the published models of NpH and KLH1 reveals that they are almost indistinguishable. The shape, principal orientation and specific structural features of the wall, such as the “anchor”, are highly conserved (see Fig.17+18).

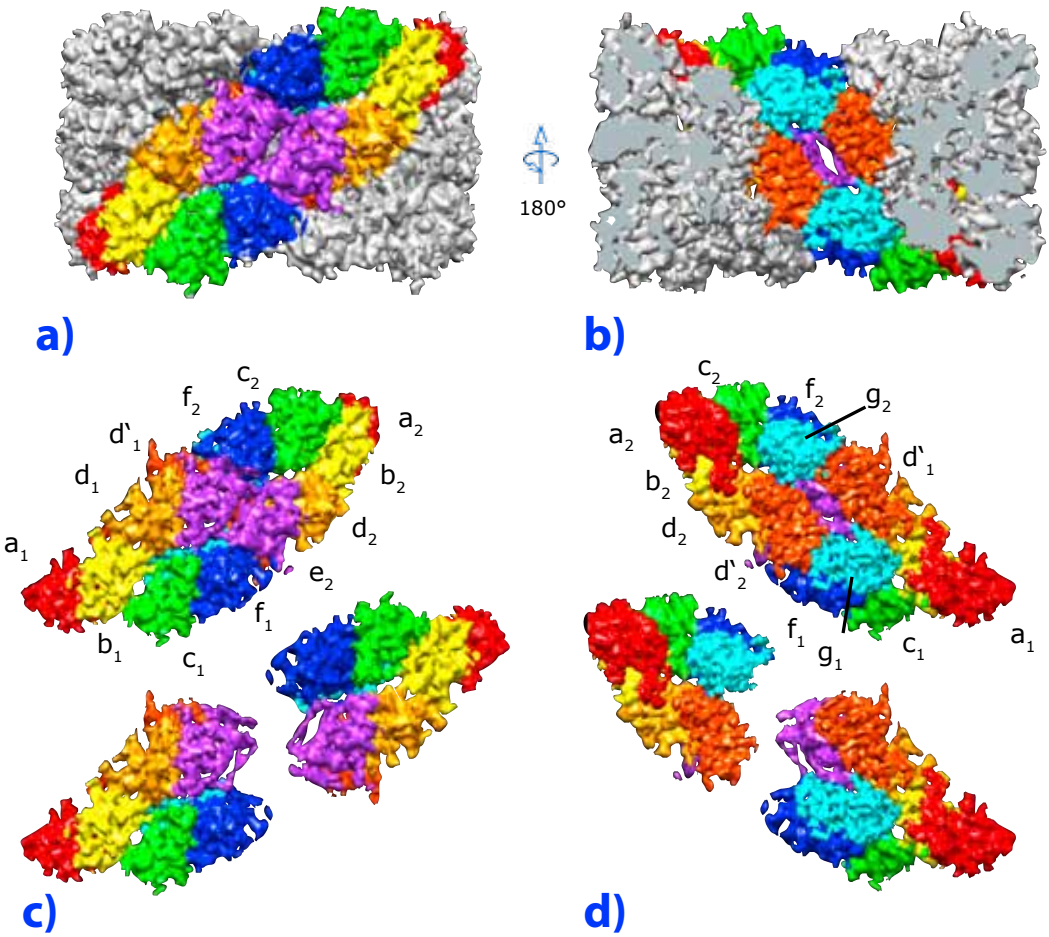


Fig. 17: Visualization of the dimer pathway and its substructures extracted from the cryo-TEM map. (a) The dimer, (b) cut open side view. Note the arrangement of the four arc FUs. (c) Extracted dimer divided into its monomers and (d) same conformers rotated by 180°. Note the classic parallelogramm view. (Color code of the FUs: a, red; b, yellow; c, green; d, orange; d', orange-red, e, purple; f, blue; used throughout)

An interesting question remains as to whether this wall architecture is fundamental to all molluscan hemocyanins, even in *Sepia*, where a gene duplication event led to a repeated FU-d (FU-d'), in the centre of the polypeptide chain. Therefore, this special case provides a good subject to prove the generalization of the wall architecture for all molluscan hemocyanins. In other words, is the hemocyanin wall still composed of the FU a-b-c-d-e and f, or is FU-d´ integrated within the wall.

The sub-nanometer resolution (8.8Å=FSC 0.5) of the model presented here enables one to

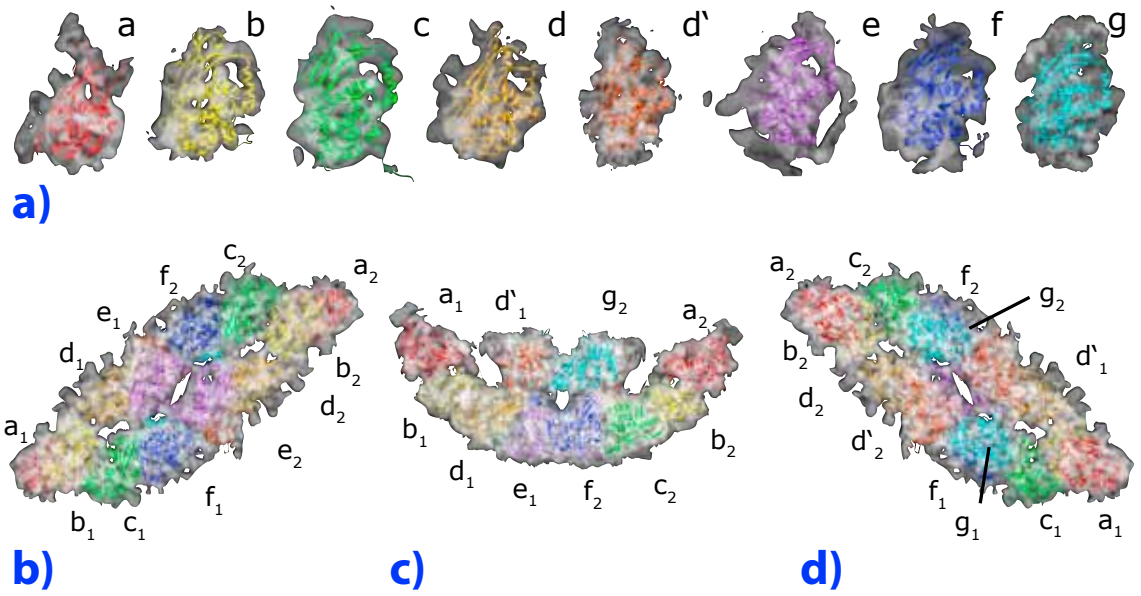


Fig. 18: (a) Extracted FUs with their respective homology models docked. Note the fitting quality (exposed helix α15 and β sheet part) and the similarities among the FUs. Extracted dimer in three characteristic views: (b) side view, (c) man-in-the-boat-view and (d) collar view.

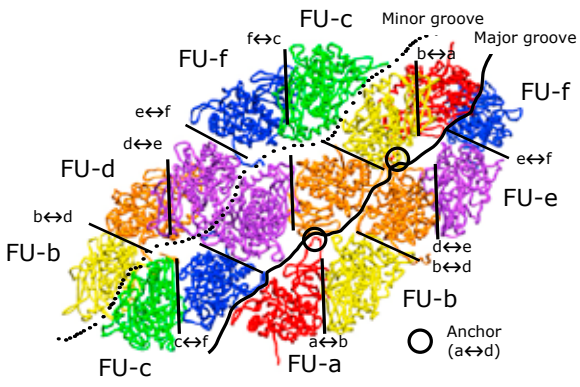


Fig. 19: Topography of the inter-FU interfaces in the hemocyanin cylinder wall. The FU arrangement is identical with NpH. The morphological unit interfaces correspond to the contact between the two protomers in the crystal structure of OdH FU-g (a↔b, c↔f and d↔e). Two horizontal interfaces connect central and peripheral tier (b↔d and e↔f). Two interfaces are across the minor groove (b↔c and d↔b) and four across the major groove (a↔f and d↔d as well as a↔e and d↔b). The circle marks the position of the anchor interface. The remaing wall/arc interfaces are shown in detail in Fig. 28 (also compare with Fig. 6 of Gatsogiannis et al. 2007).

follow the polypeptide chain and to analyze the inter-FU architecture of this hemocyanin. Therefore the standard procedure of rigid body fitting of modeled FUs and correlation of linker polypeptide length was used to determine the exact pathway of the individual

Functional unit	Rotamer-outliesrs	Ramachandran outliers (%)	Ramchandran favored (%)	C^{β} deviations
SoH1-a	3.41	0.49	97.30	3
SoH1-b	1.39	1.21	95.88	3
SoH1-c	3.35	0.73	95.86	0
SoH1-d	3.79	0.24	95.16	1
SoH1-d´	1.64	1.21	95.40	1
SoH1-e	3.49	0.95	95.23	5
SoH1-f	3.83	1.21	94.90	2
SoH1-g	2.82	0.98	95.59	7
OdH-g	7.38	0.00	97.49	0
RtH-e	16.00	5.86	76.40	64
Goal	<1	<0.2	>98	0
The MolProbity software package was applied.				

Fig. 21: Evaluation of the protein geometry of the molecular models of the eight *Sepia* hemocyanin FUs, compared to the crystal structures of *Octupus* FU-g and *Rapana* FU-e. As already described for NpH (Gatsogiannis et al. 2007) the values of the SoH models and OdH-g are convincing, whereas the values of RtH-e are surprisingly poor.

Linker peptide	a-b	b-c	c-d	d-d‘	d‘-e	e-f	f-g
Number of Amino acids	19	13	20	19	29	16	17
Maximum length (3.8 Å per amino acid)	72	49	76	72	110	61	65
Potential length (2.8 Å per amino acid)	54	37	57	54	83	46	49
Distance to be bridged (Å)	34	33	32	43	78	37	26
Note that the maximum length is highly unlikely. Thus 75% of the maximum was assumed as potential length. The minimum distance to be bridged means the distance between strand β13 of the precceding FU to strand β1 of the following.							

Fig. 20: Available linker lengths of the inter FU-linkers. Note the exceptionally long linker between FU-d and FU-d'! It is obvious, that this linker is the only one which can bridge the distance between these two FUs.

subunit (see Fig.18,19+21) (Gatsogiannis et al. 2007; Gatsogiannis and Markl 2008). In the SoH work it was reliably thought that FUs a-b-c and d will have the same position as in other molluscan hemocyanins. These FUs should not have changed their relative position, because at this point no difference at the polypeptide chain exists. However, the structural appearance of the FUs and the anchor structure between FUs a-b and d is very close to that of the other hemocyanin structures currently available. This strongly suggests that any difference would only occur after FU-d. Furthermore, it is obligate that each FU-type is represented only once in a subunit. Again the measurement of potential and actual linker length, as well as densities between the FUs (through which

Interface	SoH isoform-1	SoH isoform-2
a↔b (wall morphological	FU-a: H59, R306, W351, Y355, P356 FU-b: H483, W770, R771, F772, D773,	FU-a: H61, R308, W353, Y355, P358, FU-b: H481, W772, R773, F774, D775,
c↔f (wall morphological unit interface)	FU-c: H895, E1147, H1155, W1184, D1187,R1188,A1189 FU-f: H2557, W2842, H2843, H2807, D2845, R2846	FU-c: H895, E1147, H1155, W1184, D1187,R1188,A1189 FU-f: H2557, W2842, H2843, H2807, D2845, R2846
b↔d (horizontal tier interface)	FU-b: R444, D448, K449, N452, K453, H490 FU-d: R1256, Q1277, H1320, E1329, R1334, R1365, R1371	FU-b: R444, D448, K449, N452, K453, H490 FU-d: R1256, Q1277, H1320, E1329, R1334, R1365, R1371
e↔f (horizontal tier interface)	FU-e: H2102, T2105, R2106, K2109, D2114, L2118, H2125, Y2339, L2341 FU-f: E2510, R2511, D2512, Q2514, R2522, H2566, R2579, H2580	FU-e: H2106, R2110, K2113, D2119, S2122, H2129, Y2343, L2345 FU-f: T2514, R2515, D2516, Q2518, R2526, H2570, R2583, H2584
b↔c (minor groove interface)	FU-b: 417DVTK420, S429, D432, 511PY-WDW515, D514, Y535, E569, V572, D574, F629, W630, H632, H633 FU-c: H651; E848, R853, K854, 932PYWDW936, S984, F983, H993, F1040,Y1068, D1115, E1118	FU-b: 419DTSH422, T431, T435, 513PY-WDW517, D516, Y537, Q571, D575, F631, R633, H634, H635 FU-c: , H651 E848, R853, K854, 932PYWDW936, S984, F983, H993, F1040,Y1068, D1115, E1118
e↔e (minor groove interface)	FU-e: R2091, Q2095, E2096, 2173PYW-DW2177, F2291, H2294, H2295, F2311, M2367, S2374, Q2377, Q2384	FU-e: R2095, Q2099, E2100, 2177PYW-DW2181, F2295, H2298, H2299, F2315, M2371, S2378, Q2381, Q2388
a↔f (major groove interface)	FU-a: H106, 108LFID111, A117, Y123, 135ARAVD139, 141RL142, H178, H182, 188RFTH191 FU-f: D2606, 2607YYDV2610, 2613DS-VIENP2619, 2631FTVR2634, I2636, H2676, H2680, 2686YQKY2688	FU-a: H107, 109LFID113, A118, Y125, 138ARAVD143, 141RL142, H180, H184, 190RFTH194 FU-f: A2610, 2611YYDV2614, 2617DS-VIENP2623, 2635YTVR2639, I2640, H2680, H2684, 2690HQKY2694

Interface	SoH isoform-1	SoH isoform-2
d↔d (major groove	FU-d: D1354, 1357ND1358, 1361YY1362, S1364, T1366,	FU-d: R1358, 1361ND1362, 1365FY1366, S1368, T1370,
a↔e (major groove	FU-a: 385IN386, 388TL389 FU-e: T2133TAKH2136, C2141, H2152	FU-a: 387IN388, 390TL391 FU-e: T2137TAKH2150, C2145, H2156
b↔d (major groove interface)	FU-b: 471DA472, N474, C479, H490 FU-d: N1633, N1635, 1636IAS1638	FU-b: 473DA474, N476, C481, H492 FU-d: N1635, N1637, 1640IAS1644
a↔d (anchor interface)	FU-a: 325KDCRH329, D331, C333 FU-d: P1573, 1577EECNC1580, N1582	FU-a: 327KDCRq331, D333, C335 FU-d: P1576, 1581EECNE1582, N1586
d↔d´ (arc/wall interface)	FU-d: 1577EE1578, L1617 FU-f: 1885RYRKL1889, T1957	FU-d: 1582EE1583, L1622 FU-f: 1889RYRKL1884, T1963
d´↔g (arc interface)	FU-d´: 1906RPF1908, 1910NSTANHD1916 FU-g: R3230, 3272DAYD3275	FU-d´: 1910RPF1912, 1914NKSANHD1920 FU-g: R3234, 3276DAHD3279
f↔g (arc/wall interface)	FU-f: 2793RQKE2796 FU-g: D3009, S3015, L3017, 3050DD3051, H3198	FU-f: 2797KQ2798, K2801, K2803 FU-g: D3011, 3052DD3053, H3202

Fig. 22: Amino acid residues involved in interfaces between the FUs

a potential linker could pass) led to the complete topology of the decamer. The linker length needed to bridge the central FU with the wall FU-d (in NpH and KLH1) is 78Å ((see Fig.21). The potential length (75% of maximum length) as well as the maximum length of only one linker can bridge this distance. This is the linker length of the additional FU-d´! Accordingly, it is clear that the whole wall topology persists in SoH and that the new FU-d´ is integrated in the arc (see Fig. 19). The second question is, which of the two different arc FUs is FU-d´ and which is FU-g? The solution to this question was can also be easily provided. According to their distance, the likeliest arrangement is that the FUs d´ form the center, while FUs-g are located in the periphery of the arc. These strong arguments led to the solving of the complete topology of SoH. The interfaces shown in consequence are adapted from the interfaces presented in NpH and KLH1 (Gatsogiannis et al. 2007; Gatsogiannis and Markl 2008). Since the present specimen was not purified with respect to the two SoH isoforms, previous to the preparation and single particle analysis both molecular models are shown equally to avoid a model bias (see Fig 23-28).

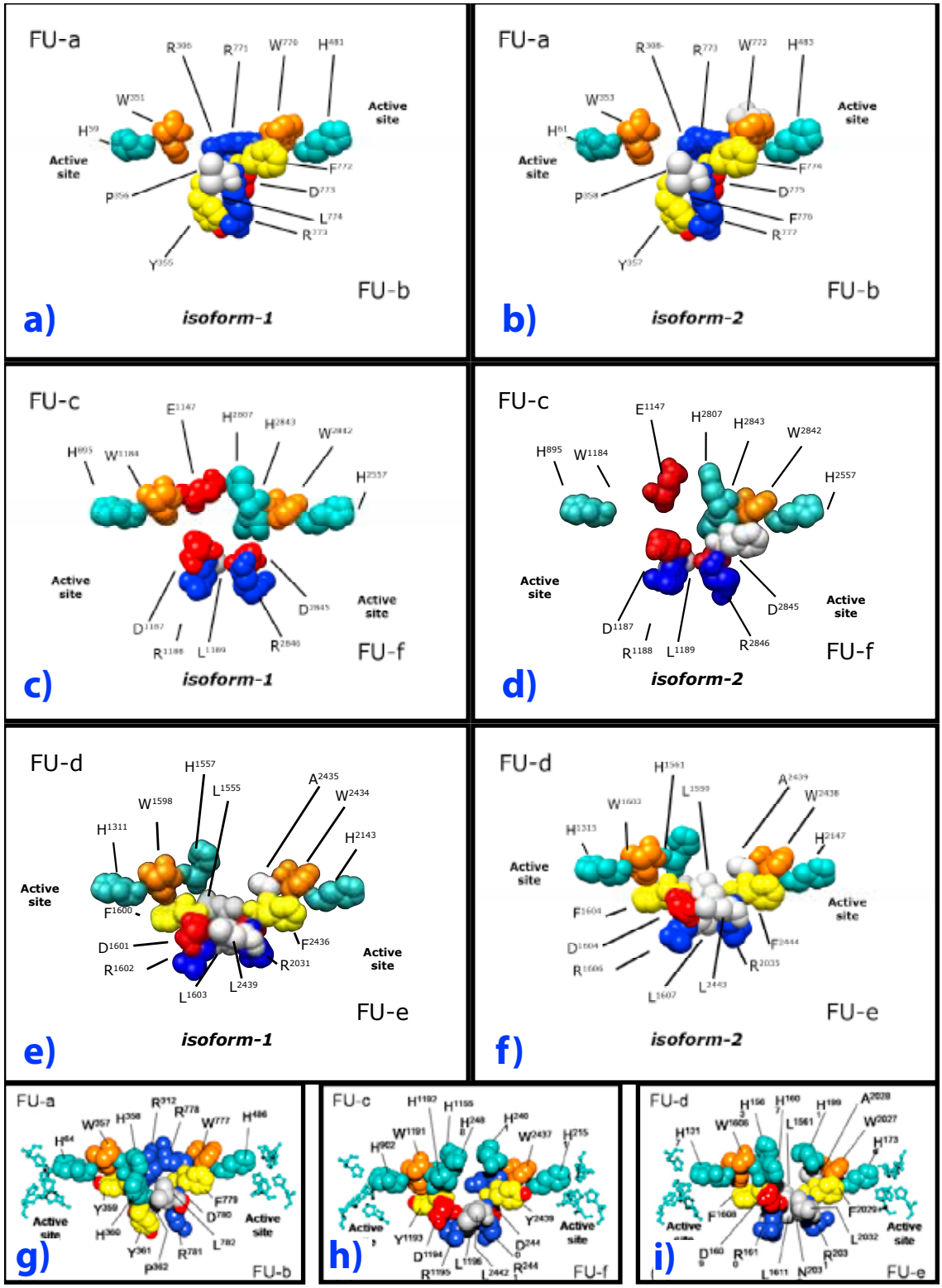


Fig. 23: Morphological unit interface arrangement in SoH (a-f) as deduced from rigid body-fitting. Amino acids shown analogue to NpH (g-i). Isoform-1 and -2 exhibit an almost identical interface architecture, while the comparison with NpH shows differences and similarities. Note the close distance charged residues at the utter side of all interfaces, in SoH and NpH. The amonunt of histidines, taking part in interface composition, is significantly reduced in SoH compared to NpH. Color code of residues: red, acidic; blue, basic; purple, polar; light sea green; histidine; orange, tryptophan; grey, hydrophobic; yellow, phenylalanine; yellow/red, tyrosine; grey/yellow, cysteine and methionine.

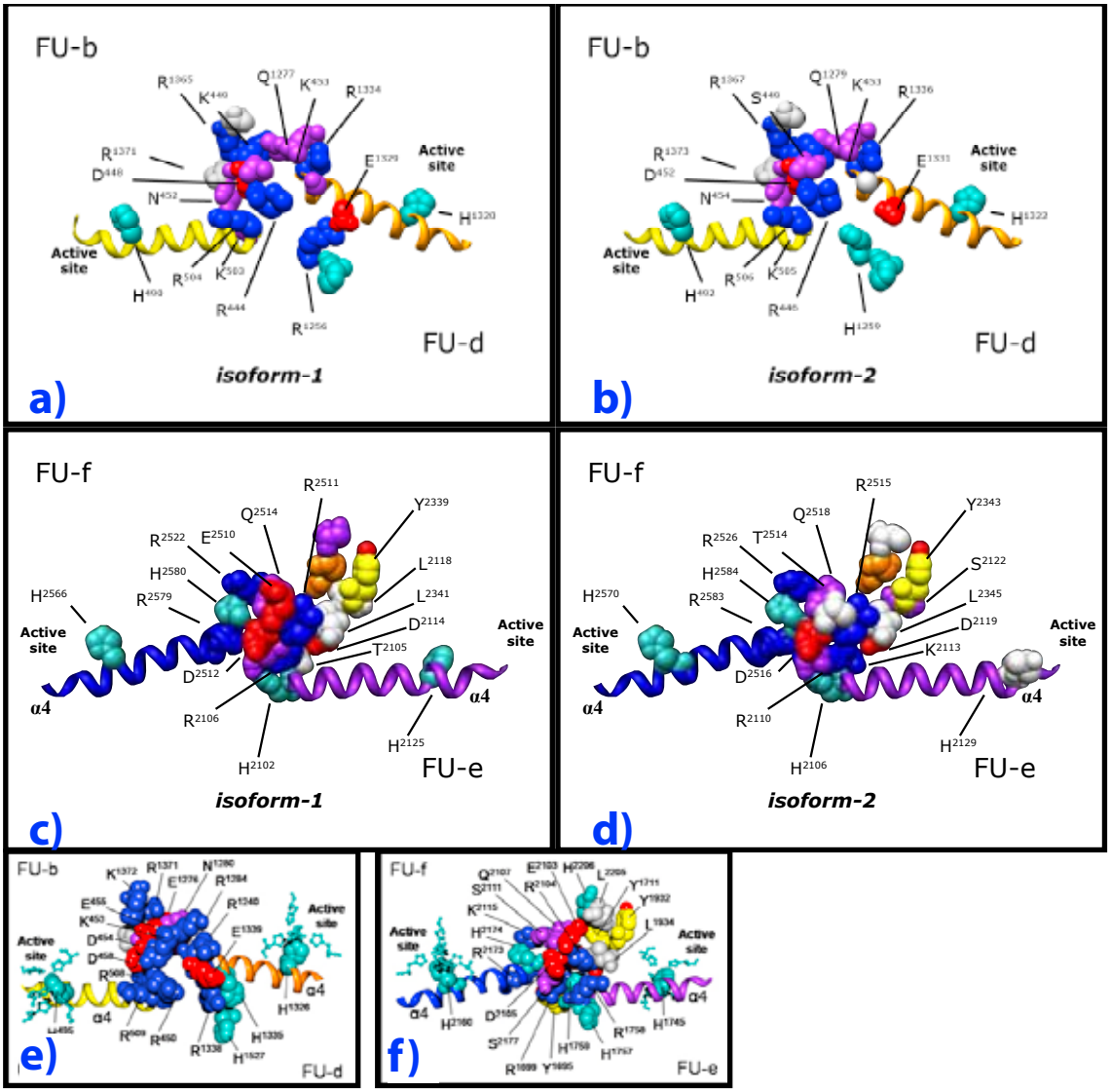


Fig. 24: Horizontal tier interfaces b↔d and f↔e (a-d) of SoH and NpH (e+f). Among the two SoH-isoforms no significant differences are observed. Although fewer charged residues take part in the interface formation compared to NpH (e+f), still a significant amount is present in SoH. (Color code see Fig. 23)

The interfaces a↔b, c↔f and d↔e (morphological unit) (see Fig.23) of both isoforms show almost no difference. However, compared with the model of NpH some differences can be measured. The central histindes are lost in many cases. In FU-c H1155 became a glutamtate (E1147). In FU-a H360 and H358 are lost: the same is true for the histidines in the core of d↔e. This is interesting because these histidines are thought to mediate allosteric signals from one active site towards another in NpH (Gatsogiannis et al. 2007). It is also significant, that the charged amino acid residues, mainly located at the bottom of the interface, are all conserved. A similar situation can be observed at b↔d and f↔e (see Fig.24). Very few changes are found to occur when comparing the two SoH isoforms.

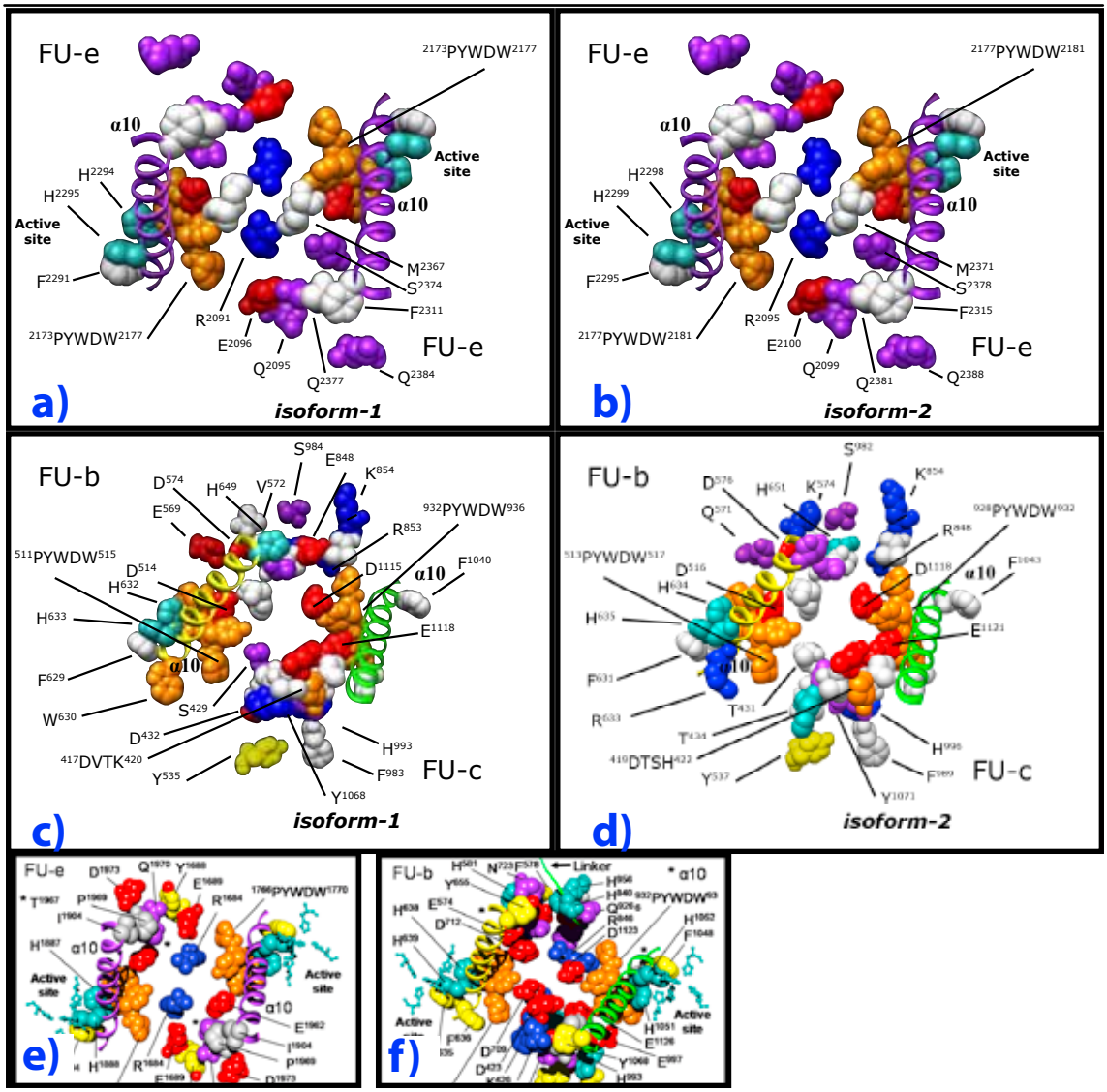


Fig. 25: Minor groove interfaces e↔e and b↔c. The comparison shows no differences between the two isoforms of SoH in e↔e (a+b) and only slight changes in b↔c (c+d). Also the interfaces of NpH (e+f) are very similar. Additionally to the highly conserved PYWDW motif (orange), opportunities for salt bridges are alike between NpH and SoH. (Color code see Fig. 23)

The major changes between NpH and SoH are not present in the charged cluster of amino-acids, but small differences in the periphery of this interface. In b↔c and e↔e the highly conserved sequence PYWDW does of course persist (see Fig.25). Again, the principle architecture is strikingly similar in both SoH isoforms. Differences between NpH and SoH can be found in the periphery mainly. The charged residues in the center of this interface are conserved. This plot continues over the remaining interfaces f↔a, d↔d and the anchor structure (see Fig.26).

In summary it can be said that the differences between SoH isoform1 and isoform2 are so minimal, that it is highly unlikely that a different structure would be detected, due

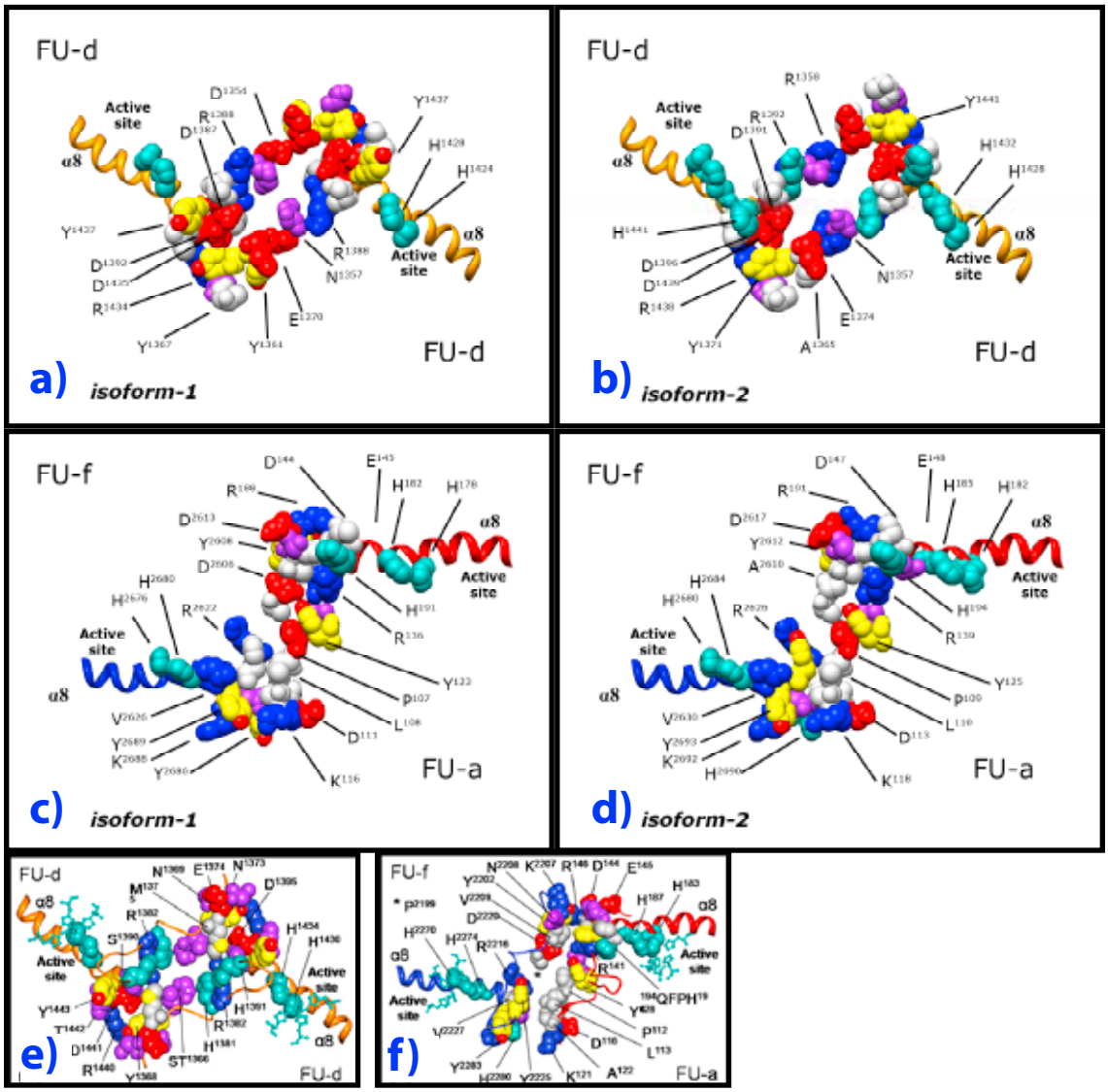


Fig. 26: Major groove interface arrangement of SoH (a-d) compared with NpH (e+f). The two isoforms show small changes around the central cavity, while the amino acid composition remains. These areas provide several opportunities for chemical bonding and active site interaction over helix α8. The situation in NpH is very similar, although some differences can be found at the central part of these interfaces. (Color code see Fig. 23)

to the isoform heterogeneity. The comparison with NpH shows that the charged amino acids in the center of the FU interfaces persist, while differences mainly appear in the periphery. This is interesting as a possible means for allosteric signal transfer. Signals are transmitted from the active site across the interface over helix movement (Gatsogiannis et al. 2007). It is thought that in NpH the histidines in the core of the interfaces a↔b, c↔f and d↔e play a significant role, together with other charged residues. Interestingly several of these histidines are absent in SoH FU interface organization, but the charged residues all remain. One can now speculate that the main allosteric signal transfer is transmitted between the FU interfaces through salt bridge rearrangements, because the

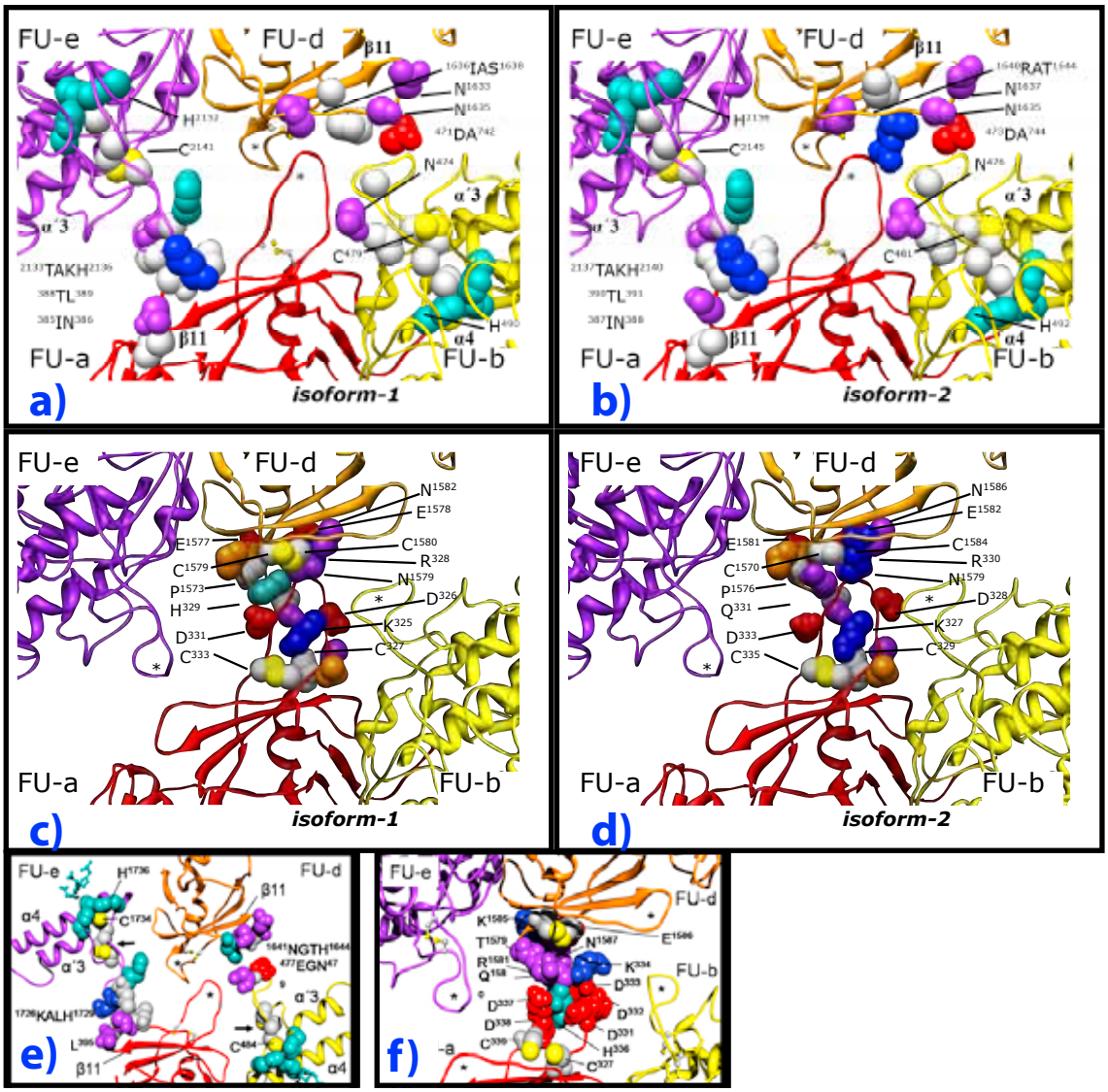


Fig. 27: The anchor interface arrangement in SoH1 and SoH2 (a-d) and NpH (e+f). Interface a↔e and d↔b are almost identical in all three models (a,b,e). The thioether (C2141/2145 and C479/ 481) bridges, involving active site histidines, are exhibited in all three cases (arrows). Interface a↔e (anchor) is highly similar in both isoforms (c+d) but shows differences to NpH (f). The aspartatic cluster of NpH is not represented in SoH. The cystidine residues forming disulphide bridges are exhibited in all three cases. (Color code see Fig. 23)

structures needed for this are conserved.

After NpH and KLH1 (Gatsogiannis et al. 2007; Gatsogiannis and Markl 2008) SoH is the third *pseudo* atomic resolution molecular model of a molluscan hemocyanin. Taken together, all three reconstructions provide an excellent overview of how this group of giant molecules is designed. The molecular models revealed 12 different types of wall interfaces. In the preceding publications already mentioned, they have been described in terms of possible contribution to allosterism and mechano-chemical interaction

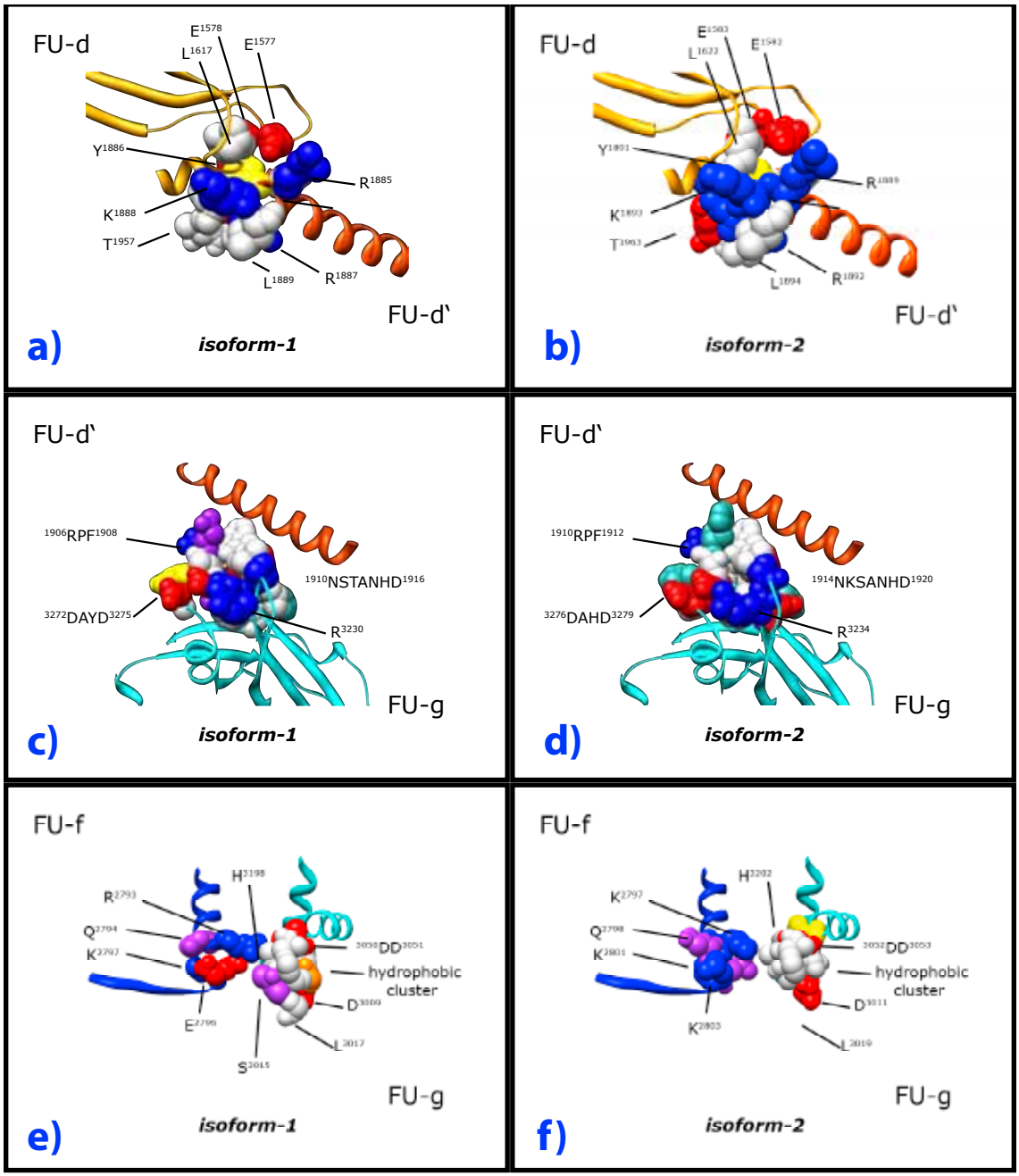


Fig. 28: Architecture of the wall/arc connection in SoH. Interfaces d↔d' (a+b), inter-arc-interface d'↔g (c+d) and f↔g (e+f). Due to the unique collar architecture of SoH these interfaces were not described, yet. In interface d↔d' a hydrophobic pocket with residues of opposing charge inside is exhibited. Note that the negative charged residues (E1557+E1578) belong to FU-d, while the positive (R1885, R1887 and K1888) belong to FU-d'. The comparison of SoH1 and SoH2 shows high similarities. However in SoH2 even more charged residues are present. The interface between the two arc-forming FUs (d'↔g) is a compsition of clustered, charged amino acids. Again SoH2 even posseses more charged residues and additional histidines. Interface f↔g is composed of a hydrophobic cluster (FU-g) and provides opportunities for salt bridges. FU-g supports this through three negative residues (D3009, D3050 and D3051) and FU-f with two, respectivley three postitve resiudues (R2793, K2795). In SoH2 E2796 is interchanged for a positve amino-acid (K2803) which could promote even stronger binding. (Color code see Fig. 23)

(Gatsogiannis et al. 2007). This work again showed that the principle architecture of these 12 interfaces is fundamental to all molluscan hemocyanins investigated so far.

The new interfaces

Three new types of interfaces are exhibited by the SoH D5 symmetric decamer (see Fig.28). The additional FU-d' is positioned in the inside of the hollow cylinder forming the arc and leaving the FU topology of the wall unchanged. All of the new interfaces offer similar possibilities for inter-FU signal transfer, as already described for the all the other interfaces. The wall-collar connection d↔d' exhibits a hydrophobic pocket containing positively charged residues (R1885,R1887,K1888) from FU-d' and negatively charged (E1557,E1578) from the opposing FU-d. This gives the possibility for salt bridges. Additionally, the covalent linkage from FU-d to FU-d' certainly enforces the bond between the FUs. The arrangement of alpha xy with the beta sheet part of FU-d allows a "push-pull" mechanism between the two FUs. This could indicate changes in spatial orientation. The second wall collar connection between FUs-f↔g is similar. Three negative residues (D3050, D3051, D3009) embedded in a hydrophobic pocket oppose the positively charged (K2795, R2793) in FU-f. Again, a covalent linker supports this connection. The third new interface is a contact between FU-d' and FU-g. Thus, it is situated inside the cylinder. Again, this interface offers possibilities for salt bridge formation. On the left side, negative residues (D3272 and D3275) oppose an arginine (R1906). The whole interface is surrounded by a hydrophobic pocket. All in all the new interfaces are very tight and the arc structure within in the collar is far more densely packed with FUs compared to those found in NpH. This tight package might also result in fewer conformational changes due to allosterism.

Comparison of the present SoH model with those in previous publications

Prior to the reconstruction presented here, two 3D structures of SoH had been elucidated by single particle molecular analysis. The publication from Boisset and Mouchee (2000) followed a reconstruction published six years earlier (Lambert et al. 1994). Due to the rapid development of the cryo-electron microscopy and molecular reconstruction methods lasting recent years, this first structure does not show any details when compared to the SoH model presented here, almost 15 years later. Nevertheless, the

principle architecture of the arc was uncovered and compared with *Octopus* hemocyanin. Furthermore, Lambert and his coworkers suggested several possible pathways for the monomer within the wall, which could not subsequently be confirmed (Gatsogiannis et al. 2007). A molecular feature that still persists is the anti-parallel orientation of the two SoH monomers forming the subunit dimer. Boisset and Mouchee´s work merely focused on a comparison between two datasets. Therefore, the interpretation of the data gained was relatively superficial. One of the two sets was taken by an electron microscope equipped with LaB6 electron emitter and the counterpart was taken by one with a Field emission gun. The result of this experiment was that the FEG data-set produced a significant better result and also reached a higher resolution. A comparison between the FEG reconstruction and the data in this thesis directly shows a high similarity (see Fig.29). Besides the significantly higher resolution, all key features of the wall were already visible. However, not all FUs are discernable and thus a rigid body fitting was not possible to that time. Three "pillar" structures were described as new features, of which one (P2) is an analogue of the "anchor" (see (3)). The other two "pillar" structures can be identified as contacts between FUs. The principal orientation of the FUs-d´ and g in the arc is also similar. A1 (FU-g) is arranged horizontically while A2 (FU-d´) is more or less orthogonal to it. The asterisk marks the position of the d'↔ g interface (4).

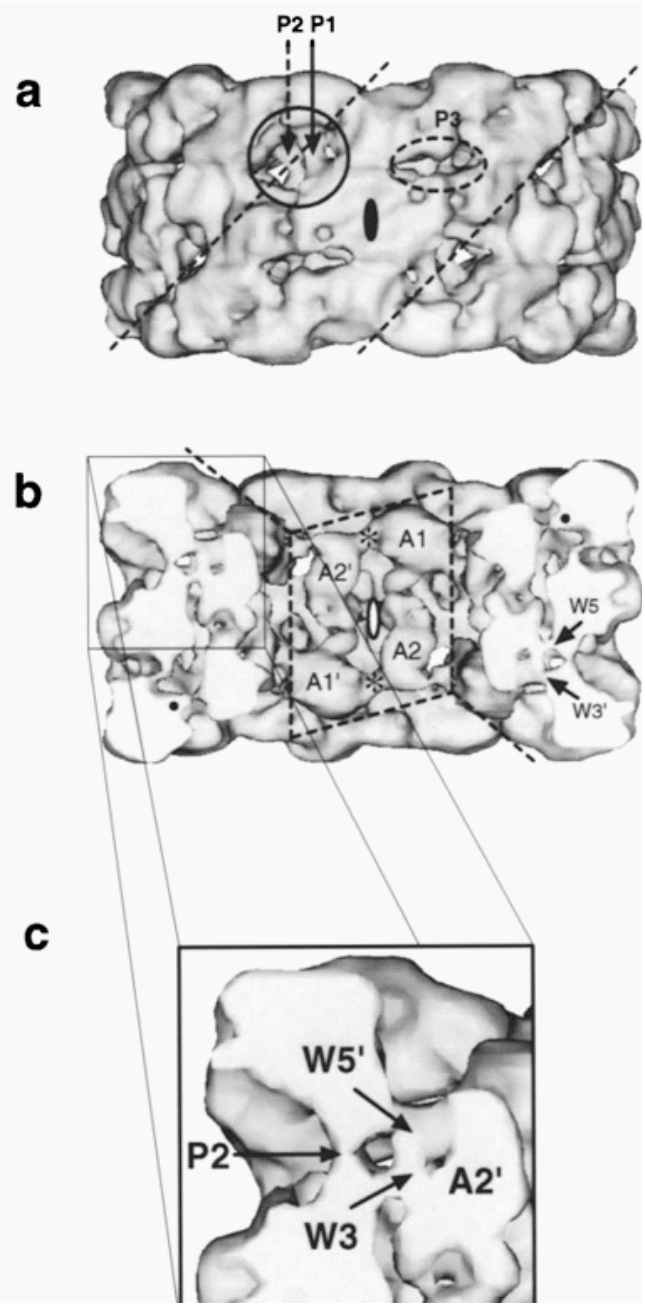


Fig. 29: Old SoH reconstruction (Boisett et al. 1999). The single FUs are discernable (FU-c especially) and the anchor structure is well pronounced (P2). The FU arrangement in the arc is identical with the arrangement presented here.

Analysis of *Nautilus* hemocyanin in oxy and deoxy-conformation

The major emphasis of this work was to use a structure-to-function approach to describe two different oxygenation-dependent conformations of *Nautilus pompilius* hemocyanin. The question was, whether it is possible to show conformational changes induced by oxygenation of the respiratory protein at a sub-nanometer level using cryo-TEM and single particle analysis. The resulting 3D-density map can then be upgraded by docking remodeled X-ray data into it. This hybrid approach should then allow visualization of interactions between the FUs in great detail. This structure-to-function approach should then give some insight into the mechanisms which control cooperative oxygen binding behavior.

The 3D-comparison of different conformations of a protein can be a very difficult task. Thus, the strategy needs to be very clear right from the start. Since the investigated molecules have a very large mass (3.5MDa), cryo-TEM seems likely to be the best method. An approach using X-ray crystallography would provide atomic resolution, but also faces severe difficulties. Crystallizing a protein of that size is very difficult and the interpretation of the enormous amount of data gained from such experiments is likely to be extremely time-consuming. A TEM negative staining preparation would be unlikely to succeed for reasons of staining artifacts, low resolution caused by the staining agent and difficulties when controlling the atmospheric composition during the preparation. Therefore the single particles cryo-TEM approach is most straightforward and can be performed within an economical amount of work. To obtain different conformations of the hemocyanin, two different assays can be devised. One approach could be the *in silico* separation of the dataset into subgroups, representing different functional states of the molecule. Such investigations have been performed very successfully with the ribosome (Klaholz et al. 2004; Penczek et al. 2006; Zhang et al. 2008; Scheres et al. 2009), GroEL in ATP bound and unliganded state as well as in CCT (Llorca et al. 1999; Ranson et al. 2001; Roseman et al. 2001). All of these experiments focus on a subdivision of the dataset into homogeneous groups using different algorithms. These experiments give the ability to observe the protein in a type of *time resolved* way, in such that intermediate steps, if existing, are also visualized. The extreme conformations of the molecule are also represented. For the hemocyanin this would mean a structure representation of a *fully* oxy- and deoxygenated state. However, this method also has dramatic drawbacks. In order to achieve such a stepwise molecular distribution into subpopulations, the user needs many thousands or even hundred thousands particles

to maintain a significant signal throughout all subpopulations. In the case of NpH, an average of 70 particles was picked from one micrograph, because of the size of the molecule. This means an incredibly large number of micrographs would need to be taken at the electron microscope to conduct the experiment. Thus, with our technical setup, this alone is an exclusion criterion for such a molecular subdivision experiment. But further complications would also occur. In a normal vitrified specimen preparation the surrounding atmosphere in the blotting chamber will have an oxygenating effect on the specimen. Accordingly, the vast majority of proteins will be in an oxy state and it would be difficult to find a significant sub-population of deoxygenated proteins within the dataset. To overcome this problem the number of images would have to be increased again. To create such enormous datasets computer-controlled automated data acquisition could be used. Such a setup would run several consecutive days recording thousands of images on a digital camera (Potter et al. 1999; Carragher et al. 2000; Suloway et al. 2005). This setup could solve the problem of data acquisition. Right now it is still very expensive and difficult to perform. However, in the context of the present work, automation was not an option, due to limited microscope time and the availability of only a low resolution digital camera. Nevertheless, further significant difficulties arise in evaluating the dataset. Maximum likelihood or bootstrap algorithms (Penczek et al. 2006; Scheres and Carazo 2009) should be able to perform a subdivision. However, the computer power available when this project started in 2006, was not sufficient to evaluate datasets with this size containing molecules of the size of NpH to a sub-nanometer level. This experiment could be worth reconsidering when improved computing power becomes available. Methods like 3D MSA (Klaholz et al. 2004) are very powerful if the data separation is founded on molecules in ligand-bound or ligand-free state, for example. Furthermore, it is very useful if the study can concentrate on a molecular domain where the structure changes. In this initial experiment no such information was given *a priori*. A distribution on a correlation-based algorithm and 2D-MSA (Elad et al. 2008) would face similar problems as the 3D MSA. Finally, at the end of all of these experiments no one could not determine which extreme group of molecules is the oxy and deoxy state, or vice versa. The 2nd possibility is to try to prepare the molecule in such way that the dataset contains only one oxygenation state of the hemocyanin. In this case two preparations are required, one with 100% oxy- and the other with 100% deoxygenated hemocyanin. Inducing the deoxy conformation by chemical release of the coppers at the active site, is not an option, since the result would not reflect a native conformational change. Thus, strict incubation of the protein specimen in a defined atmosphere containing no O_2 is the best option to

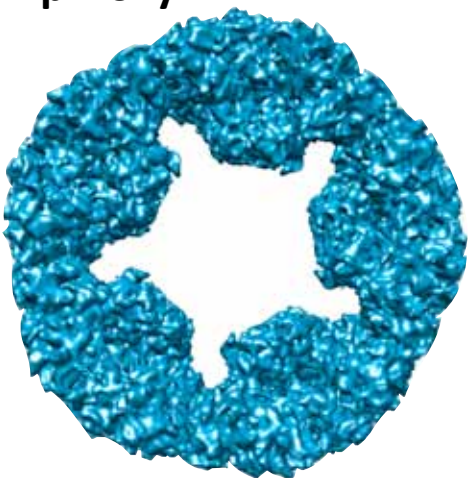
accomplish this goal. However, this approach cannot provide a time-resolved resolution of the progressive oxygen binding. Additionally, it is very likely that the sample will not be homogenous, but a mixture of many states with a strong tendency towards the desired state of the molecule.

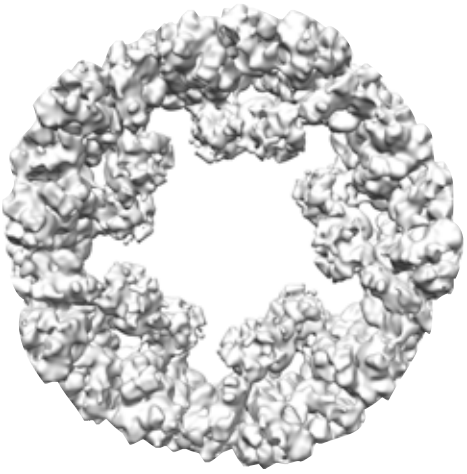
All in all, this way to prepare two molecular datasets seems to be most straightforward and best, viewed also from economical aspects. To achieve this experiment a newly designed plunge-freeze apparatus had to be designed and constructed (see Materials & Methods), capable of controlled atmospheric incubation and cryo-fixation of the specimen.

Attempts to visualize conformational changes of hemocyanins have been already performed using other methods. Previously, Hartmann and colleagues claimed to observe structural changes caused by the presence or absence of molecular oxygen, using low resolution small angle X-ray scattering techniques (SAXS), with Keyhole-Limpet-Hemocyanin (KLH)(Hartmann et al. 2003). KLH1 is one of the two isoforms of the didecameric hemocyanin of the Giant Keyhole Limpet *Megathura crenulata*. It has 160 O_2 binding sites and, along with other didecameric molluscan hemocyanins, with a mass of ~ 9MDa is one of the largest mass functionally active proteins known. Oxygen binding measurements have shown that KLH exhibits little cooperativity in comparison to the hemocyanins of other molluscs (Zolla et al. 1978; Senozan et al. 1981). Generally, the cooperative oxygen binding is very low in all molluscan hemocyanins, and didecameric hemocyanins do not significantly differ from the single decamers (van Holde et al. 1995). However, the SAXS data presented by Hartmann and collegues showed significant structural changes occurring due to oxygen binding even at low resolution. Their comparison indicated that oxy-KLH1 is more compact than its deoxy counterpart, *ie* a massive structural change. It was further shown that in comparison with the cryo EM map of KLH1 (provided by (Orlova et al. 1997a) the volume is 12% smaller in solution than in the ice. This led to the statement that the hemocyanin releases the bound oxygen upon exposure to the high vacuum in the TEM and change is conformation towards the deoxy state. Thus, all single particle reconstructions would present the deoxy state of the hemocyanin.

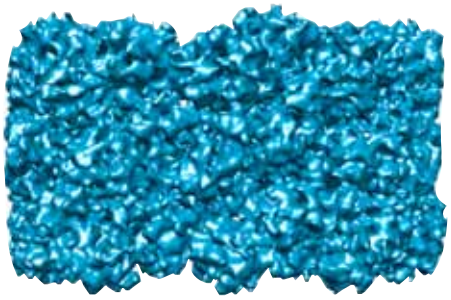
In addition to this low resolution small angle X-ray scattering data, two high resolution X-ray structures of individual functional units from molluscan hemocyanins are available. They present isolated FUs in oxy and deoxy state at atomic resolution (Cuff et al. 1998; Perbrandt et al. 2003). These crystal structures are essential for the cryo-TEM structural

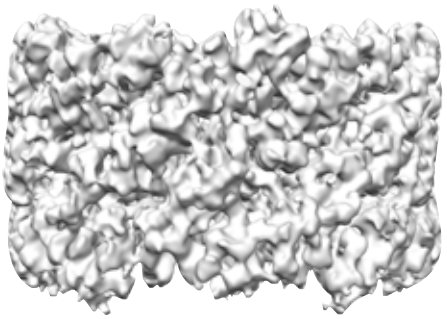
NpH oxy **NpH deoxy**



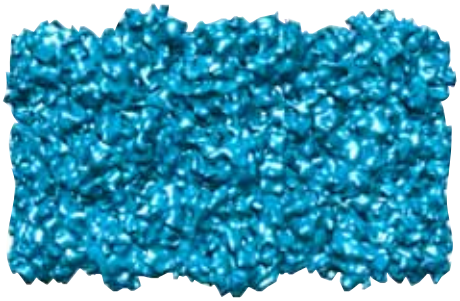


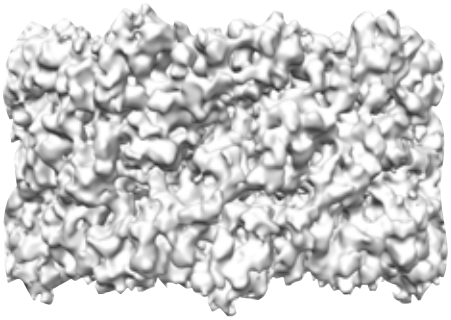
top view



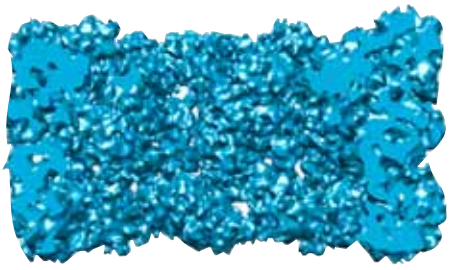


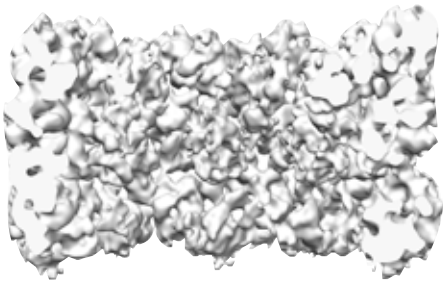
side view





side view (36°)





collar view

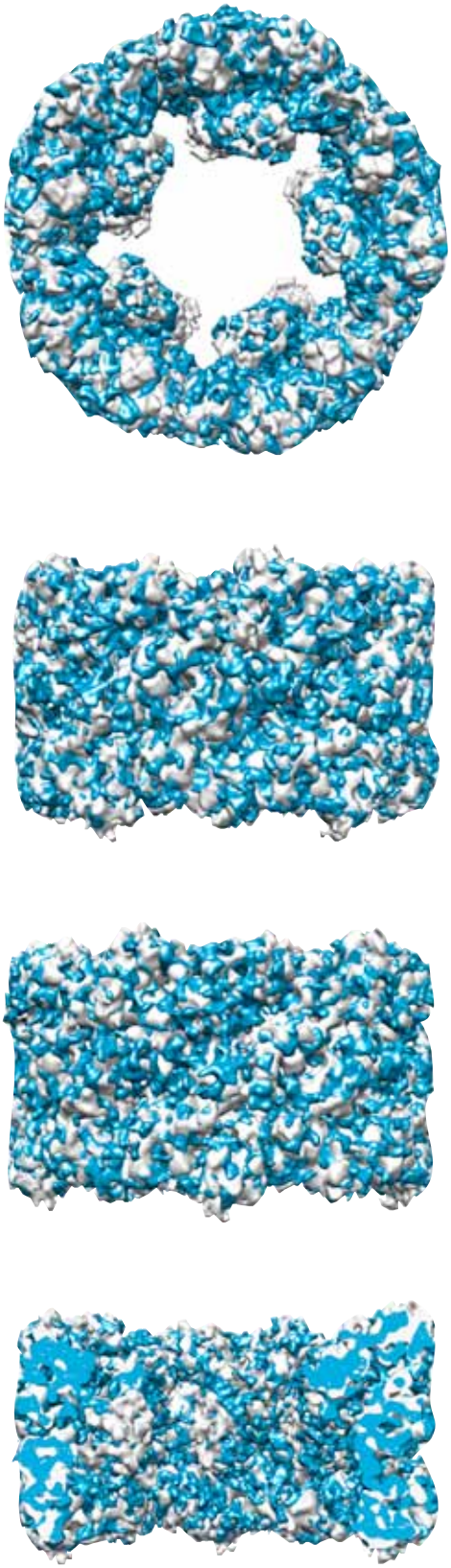


Fig. 30: The two reconstructions of NpH and a superpositon of both in characteristic orientations (oxystate blue, deoxystate grey). All FUs are well defined. Note the different densities within the FU interfaces (right page). Also note the strong movement of the arc structure (top-view and collar view). The two FUs are close together in the oxy- and are moved apart in deoxy conformation. The movement is stronger in FU-g_2 (the FU in the center). Besides this it is very difficult to directly locate areas of reorientation, still the strong differences between both reconstructions are obvious.

investigations on all molluscan hemocyanins. Together with the respective sequence, they are the basis for the molecular models of NpH and SoH (see Results I,II). Furthermore, the organization of the two FUs of *Octopus* as a morphological unit significantly helped the interpretation of the overall topology. The authors already indicated possible inter FU signaling through which allosterism could be mediated. Additionally, the comparison of the oxy and deoxy structure provides further information. During oxygenation the copper ions move 0.7Å apart (from 2.8 to 3.5Å) and the six active site histidines show a compensatory movement that might transfer forces to the four central α-helices that carry them. Thus little changes of the helix position could affect the oxygen binding properties at the active site (Gatsogiannis et al. 2007).

Very recently the work of Cong and colleagues showed structural changes in an arthropodan hemocyanin caused by SDS (Cong et al. 2009). A double reconstruction comparison approach, similar to the approach used in this work, showed that it is possible to reveal structural changes of hemocyanins using cryo-TEM maps. This work was concerned with the SDS-induced conversion of an arthropod hemocyanin into an enzyme with PO activity. While the strategy followed in this experiment is very similar to my work, the arthropod hemocyanins fundamentally differ structurally from the molluscan hemocyanins. Furthermore, the conformational changes were not induced by a change in oxygenation, but with an activation induced by SDS. Nevertheless, it was shown that changes happening at the quaternary structure level can be visualized and compared using cryo-TEM and single particle analysis. The subsequent interpretation using *pseudo*-atomic models was performed on a molecular basis. This experiment indicated that it is obligate to reach a high resolution (< 1nm) in order to visualize the conformational changes that occur at the subunit- and for the gastropod hemocyanins also the FU-level.

The results presented here are two entirely independent reconstructions with sub-nanometer resolution (see Fig 30+31). In both cases several secondary structure elements are visible, thereby underlining the resolution achieved. A superposition of both reconstructions shows, that significant differences exist between the two models. Especially the spatial orientation of single FUs varies greatly, whereas the overall shape such as diameter or height does not change significantly. The molecular model (Gatsogiannis et al. 2007) of NpH was docked into the structures using rigid body fitting to gain a *pseudo*-atomic model. The analysis of the differences depends on the description of the amino acids located at the interfaces, already published (see Fig.33-39). Additionally, in this interpretation special care was taken to define the densities

bridging the gaps between adjacent FUs and the amino acids at these locations. Distance measurements between c-alpha atoms of selected amino acids have been inserted for comparison, to describe the movements.

The differences between both reconstructions, together with the relatively high resolution of both individual reconstructions are good indications that the data is valid. The models

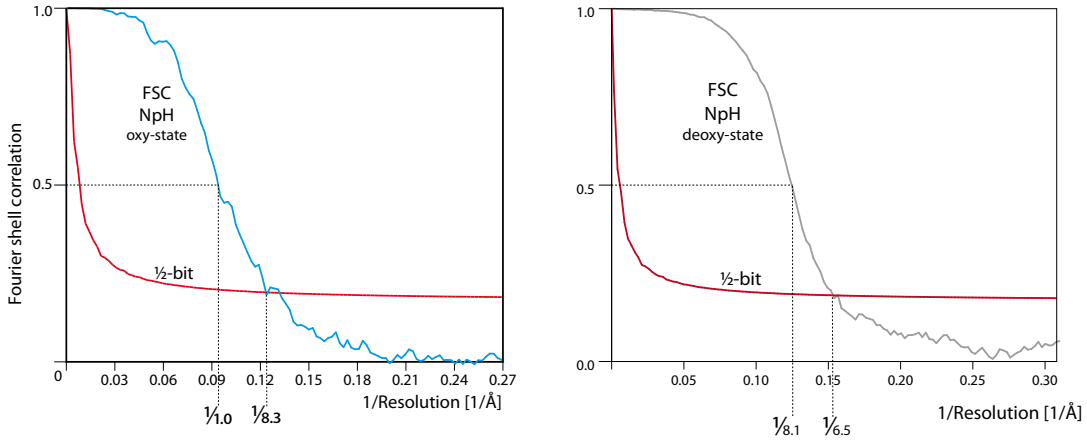
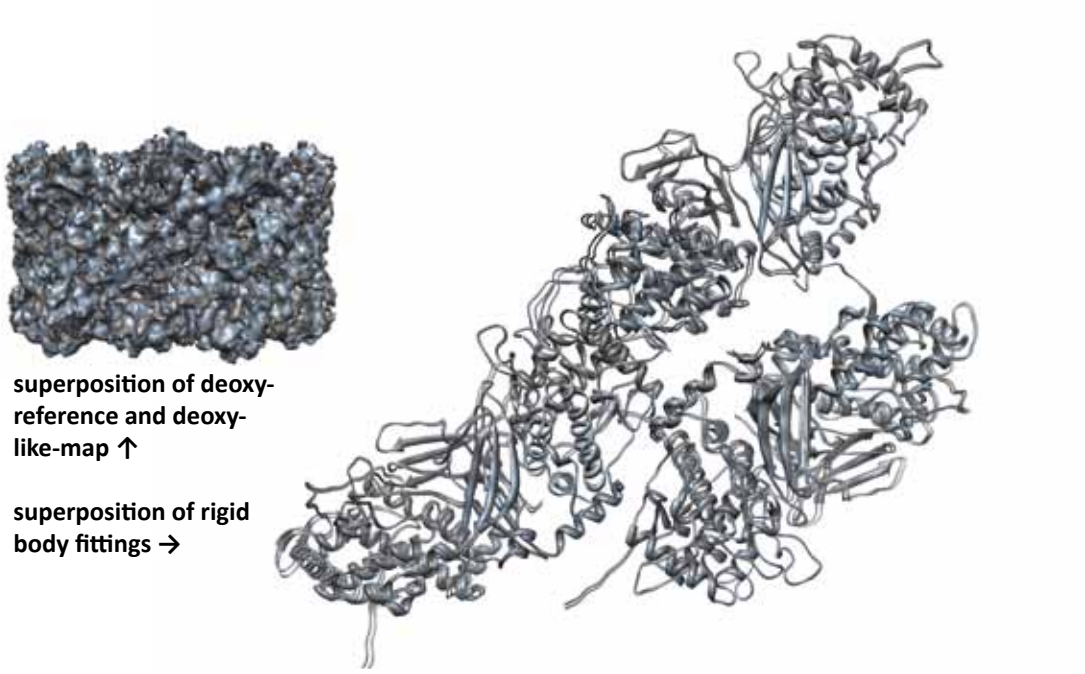


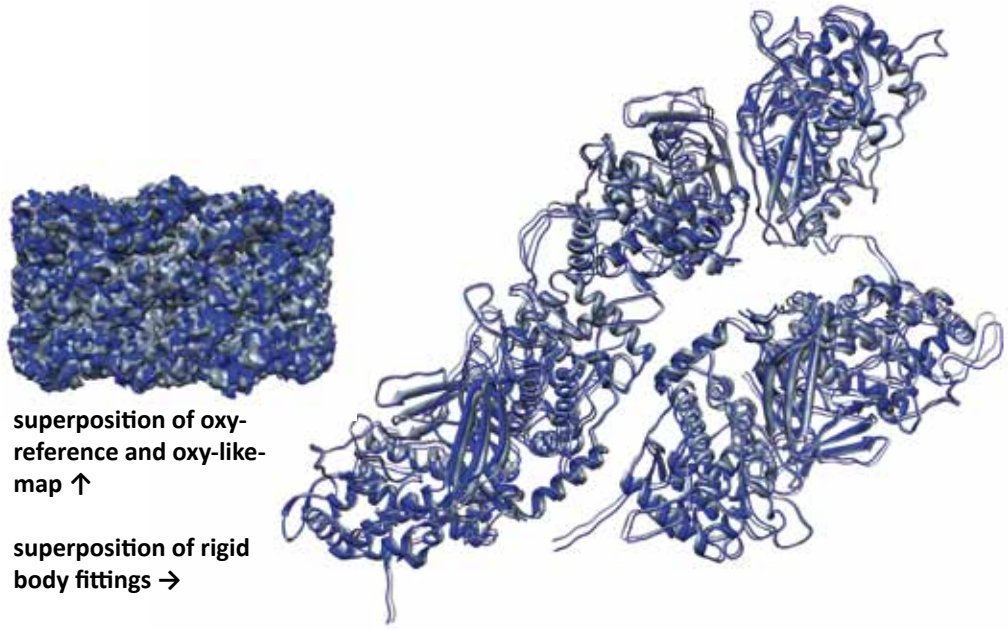
Fig. 31: The FSC-curve of both NpH reconstructions (oxy-state blue (left), deoxy grey (right)). In this work the rather conservative, but commonly used, FSC=0.5 criterion was denoted. For comparison with the publications of NpH and KLH1 (Gatsogiannis et al. 2007 and Gatsogiannis & Markl 2008) the ½-bit values are also shown (red line). (oxy map: 10Å (FSC=0.5), 8.3Å (½-bit); deoxy map: 8.1Å (FSC=0.5), 6.5Å (½-bit))

represent two conformations of the same hemocyanin. The difference should depend only on the oxygenation conditions present during the sample preparation for cryo-TEM. Consequently one oxygenated as well as one deoxygenated state has been under investigation. However, as already mentioned, it is most likely that these models are mixtures of slightly different conformations, with a strong tendency towards the desired functional state (ie complete or zero oxygenation). Nevertheless, it is not possible to prove the actual structural state of the molecule. To my knowledge no method is currently available that is capable to measure the specimen state immediately after plunge freezing of the specimen grid. Therefore, we decided to present evidence that the data does indeed accurately reflect the average of all single particles within one dataset. The final reconstructions of the datasets were used as a reference for the respective opposing dataset, to determine whether they would convert back to the correct form. This method was also used for different conformations of a chaperonin by Clare and coworkers (Clare et al. 2009). This complex experiment showed that the NpH reconstructions indeed slowly convert back to the original structural state (see Fig.32). After the final iteration, rigid body fitting was performed showing only very few variations between the reconstructions coming from the same dataset and strong differences between the datasets. As already



superposition of deoxy-reference and deoxy-like-map ↑

superposition of rigid body fittings →



superposition of oxy-reference and oxy-like-map ↑

superposition of rigid body fittings →

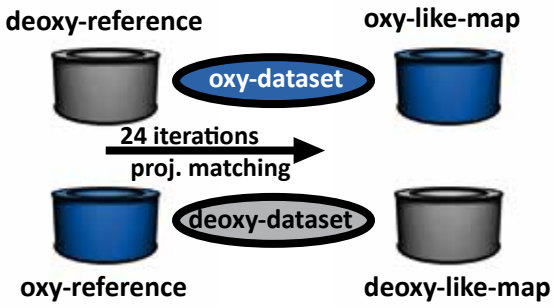


Fig. 32: Result of the cross-over experiment. After 24 iterations of projection matching the densites show high similarities. The high correlation of the ridig body fittings certainly proves that the maps converted back to the form of the dataset. (Color code: Reference maps, slate grey; oxy, deepblue; deoxy, dark grey)
←

mentioned, it has been previously suggested that hemocyanins loose their bound O_2 upon exposure to the high vacuum in the column of the TEM (Hartmann et al. 2003). This statement relies on the their claim that oxy-KLH1 in solution is more compact than deoxygenated KLH1 and that the molecules in solution are significantly smaller related to a cryo-TEM map (Orlova et al. 1997b). However the cryo-map by Orlova and collegeus was obtained using a cryo-negative staining technique (see Mat & Meth) which was not the case in the NpH work presented here. Therefore significant magnification errors could have occurred. Furthermore, such a significant dimensional change is puzzling because a careful comparison of NpH with the hemocyanin of *Haliotis tuberculata* (HtH) revealed that their diameters are almost identical, although the images of the reconstructions have been taken on different microscopes that have been calibrated by different people at different times (Gatsogiannis et al. 2007; Meissner et al. 2007). Additionally, we resampled the size of the cryo-TEM map of NpH down by 3% to show that this reduction causes many critical overlaps of the molecular models. An increase of diameter would change the inter-FU distances in such way that the interconnecting linkers could no longer bridge the gaps between the individual FUs (Gatsogiannis et al. 2007). The size relation of the two PDB FU structures deduced from X-ray diffraction shows that they are identical, although they represent different functional states of the isolated FU. Accordingly, a change of diameter could not be related to a change of single FU size. Finally, it is very unlikely that a change of size (especially increase) is possible while the protein is fixed in amorphous ice at temperatures below -150°C. Thus, it is very unlikely, if not impossible, that the structural changes observed by SAXS studies rely on different conformations due to oxygenation of the protein.

These results underline the fact that we were indeed able to accurately define two functional states of the *Nautilus* hemocyanin. Since we modified the specimen preparation in a way that it supports oxy- and deoxygenation, respectively, it is highly likely that the results presented here reflect the true oxygenation and deoxygenation conformations.

First and foremost during this study a multiple leveled signal transfer has been considered for allosteric interaction. Upon oxygenation, the hemocyanin FUs could change their spatial orientation, which in turn could cause a shift in conformation of the overall cylindrical quaternary structure. The α-helices (10,8,4), to which the active sites histidines are attached, could act as a lever and mediate the signal from the state

Interface	Amino acid residues
$g_1 \leftrightarrow g_2$ (arc morphological unit interface)	FU-g1: 2843WAFDRL2847 FU-g2: 2843WAFDRL2847
a↔b (wall morphological unit interface)	FU-a: H 64, M66, T68,240CAKELFTV247, M249,R312, R329, D379, 383YV384, 400LI401, 407DY408 FU-b: 488MA489, 596DFCD599, H663, 668EVH670, 775MTWRFDR781, 797TADDEF802, 819DII821, A823, I826, V829
d↔e (wall morphological unit interface)	FU-d: 1488AL1489, M1492, H1563, H1607, R1628, F1631, H1648, I1655 FU-e : E1911, 1914CALE1917, H1991, 2052FH2053, H2069, L2077, P2080
c↔f (wall morphological unit interface)	FU-c: H895, L1080, H1084, H1155, H1192, 1194DR1195, R1217, 1238TVL1240 FU-f: 2325AYCALE2330, H2401, H2416, H2485, 2488HE2489, F2494
b↔d (horizontal tier interface)	FU-b: R450, 453KDESID458, R508, R509 FU-d: R1262, D1272, E1276, N1280, R1284, H1335, 1338RER1340, R1371, K1372, H1527
e↔f (horizontal tier interface)	FU-e: Y1695, R1698, R1699,R1702, D1702, D1703, D1707, Y1711, 1757HTH1759, Y1932, L1934 FU-f: 2103ERD2105, Q2107, S2111, K2115, R2173, H2174, S2177, L2205, H2206
b↔c (minor groove interface)	FU-b: E425, K431, 434RRLHD438, 524MI525, L527, 529HL530, F578, 709DTLDF713, R716 FU-c: 844TRR846, I848, 851LT852, E854, 939AITH942, F991, 993HF994, Y1122, 1125LE1126, L1135, 1137RE1138
e↔e (minor groove interface)	FU-e: 1684RL1685, 1687IYE1689, R1829, E1962, M1966, Y1974
a↔f (major groove interface)	FU-a: H104, Q110, P112, L113, I115, D116, 121KAKK124, Y128, M137, 141RAVDER146, H183, H187, 193QFPH196 FU-f: E2103, H2194, P2199, D2200, Y2202, D2203, 2207KQV2209, R2216, E2220, E2223, Y2225, V2227, N2229, H2270, H2274, 2280HQRY2283
d↔d (major groove interface)	FU-d: F1355, H1357, A1360, M1362, E1364, 1367FY1368, M1375, T1377, 1379PFHR1382, I1385, 1387FI1388, 1391HTS1393, 1399ELFNS1403, 1440RD1441, Y1453
a↔e (major groove interface)	FU-a: 392NGTL395 FU-e: 1726KALH1729
d↔b (major groove interface)	FU-b: R450, 453KDESID458, R508, R509 FU-d: R1262, D1272, E1276, N1280, R1284, H1335, 1338RER1340, R1371, K1372, H1527
a↔d (anchor interface)	FU-a: C327, 331DDDKSHDDC339 FU-d: C1576, T1579, Q1581, 1586ENC1588
$g_1 \leftrightarrow d$ (arc/wall interface)	FU-g1: Q2858, N2859, R2861, H2863, H2902 FU-d: K1353, R1354, D1404, Y1407, Y1477, R1478, 1536HGLS1539, Q1542, 1544EHE1546, E1549, H1553
$d \leftrightarrow g_2$ (arc/wall interface)	FU-g2: R2712, I2716, A2719, Y2723, R2724, 2776TLEFH2780, 2782HNIPQLA2789 FU-d: C1576, 1583SSKENCN1589

Fig. 33: Amino acid residues involved in interfaces (modified after Gatsogiannis et al. 2007).

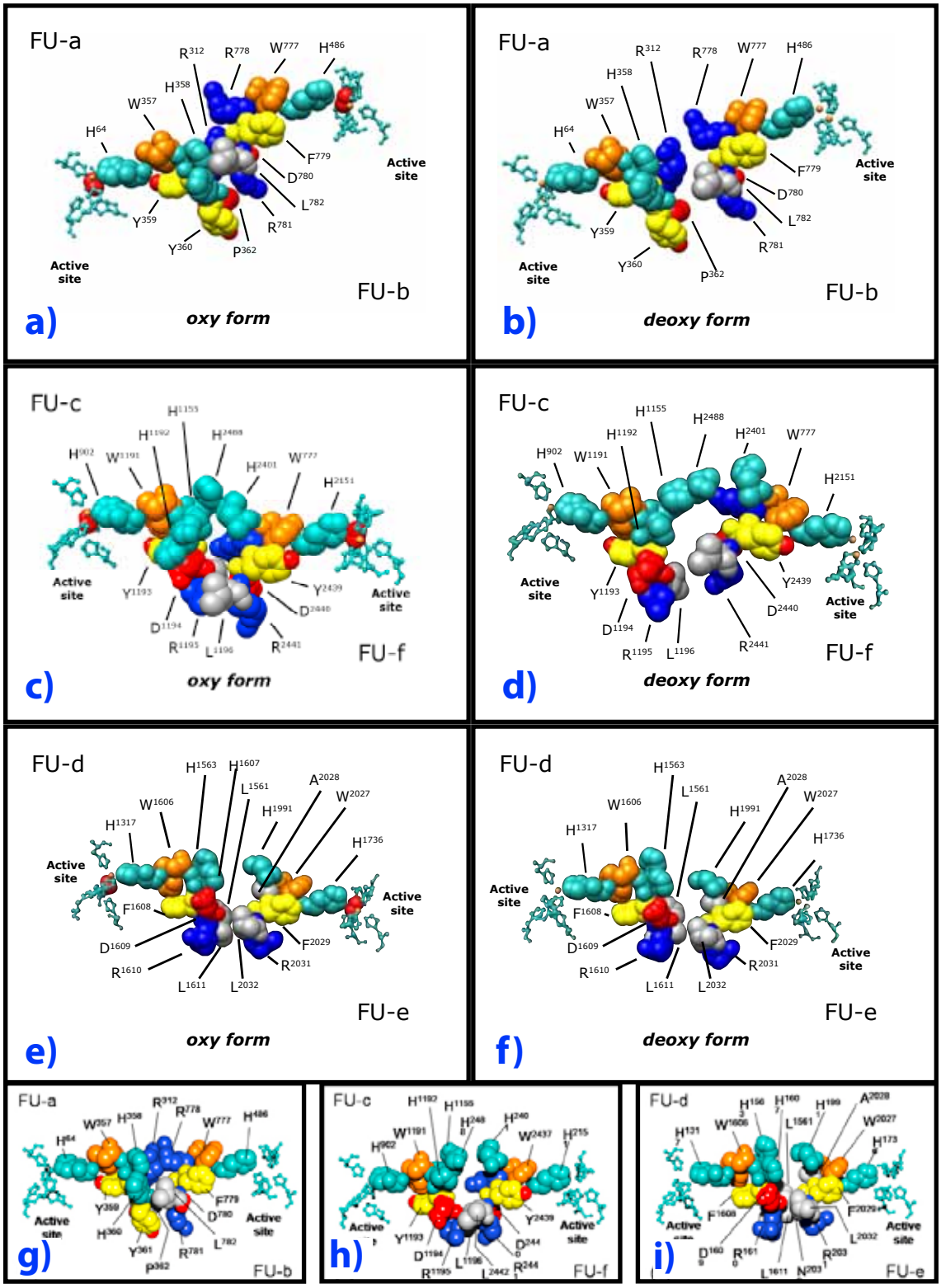


Fig. 34: Wall morphological unit interface arrangement in NpH oxy/deoxy (a-f) as deduced form rigid body-fitting. Color code of residues: red, acidic; blue, basic; purple, polar; light sea green, histidine; orange, tryptophan; grey, hydrophobic; yellow, phenylalanine; yellow/red, tyrosine; grey/yellow, cysteine and methionine. Position and orientation of interfaces shown in comparison to the published model (Gatsogiannis et al. 2007). For a detailled description see Fig. 40, 41 and 42.

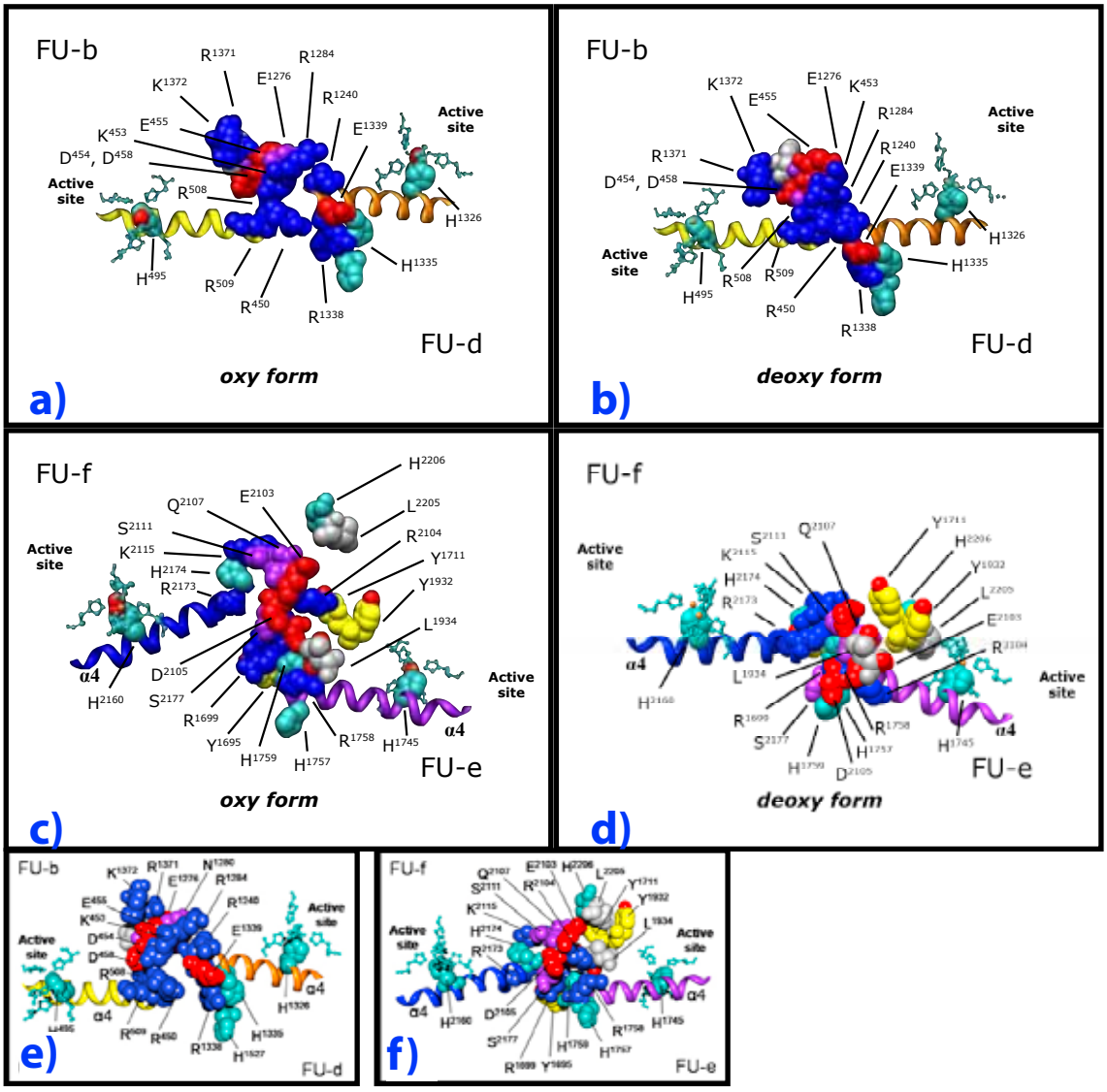


Fig. 35: Horizontal tier interfaces b↔d and f↔e (a-d) of NpH oxy/deoxy and NpH (e+f). Position and orientation of interfaces are shown in comparison to the published model (Gatsogiannis et al. 2007). (Color code see Fig. 34. For a detailled description see Fig. 43 and 46)

of the active site into the periphery, where the signal could be transferred over the interfaces towards adjacent FUs. Furthermore, ions could be bound between charged amino acids causing a mechano-chemical interaction across the interface. While all the three mechanisms are highly likely to interact the method applied here can only visualize the first directly and, if higher resolution could be achieved, possibly the second point. The movement and reorientation of single FUs is clear from direct observation of the density maps. The second point is more difficult. The extracted densities with docked PDB structures show that secondary structure elements such as alpha helices are well shaped in the periphery of one FU, but become significantly less pronounced within the FU (see Fig. 40-48). Thus, for an effort to clearly identify shifts of alpha helices which are

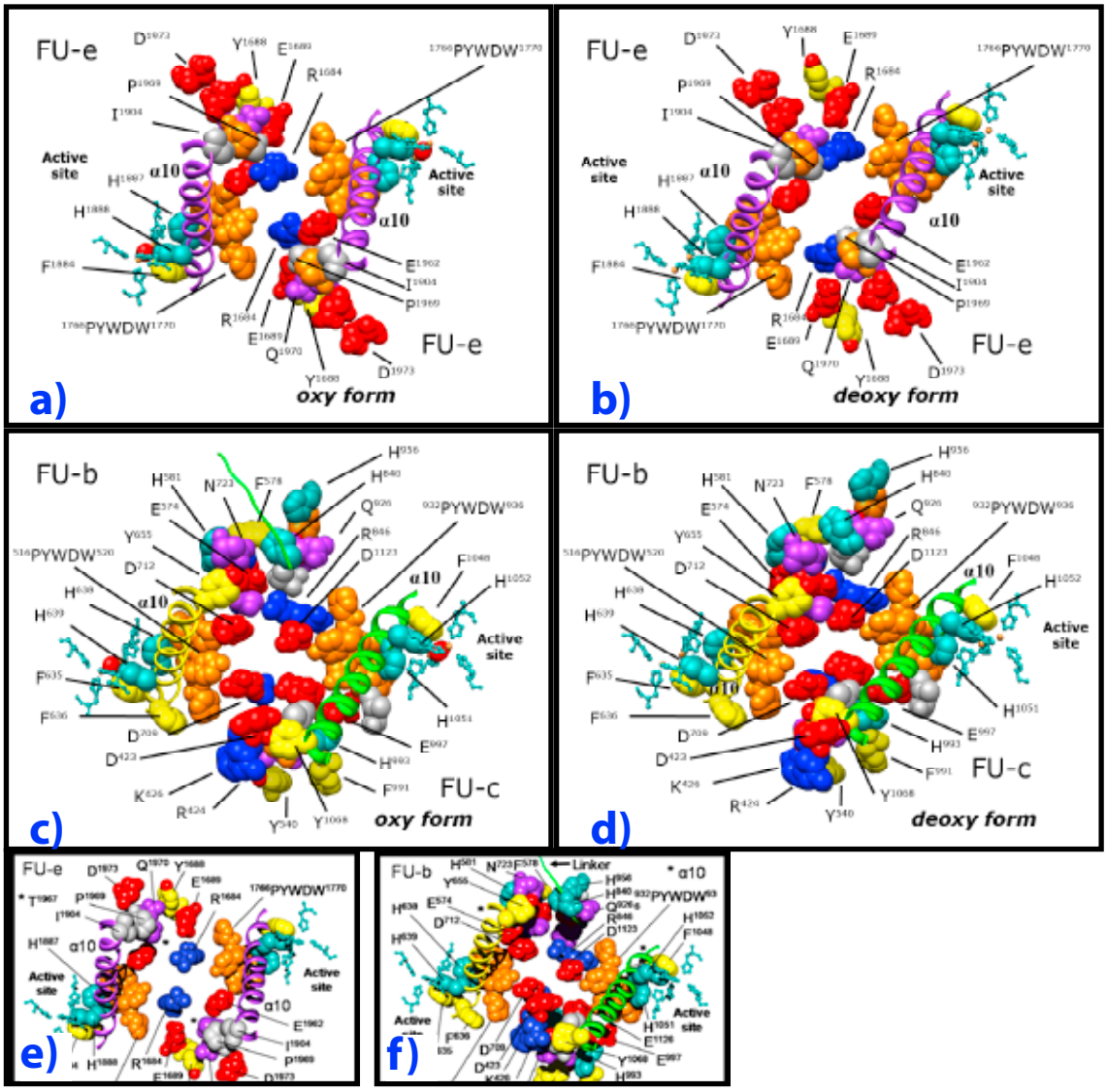


Fig. 36: Minor groove interfaces e↔e and b↔c of oxy/deoxy (a-d) conformation and published models (e+f). Position and orientation of interfaces are shown in comparison to the published model (Gatsogiannis et al. 2007). (Color code see Fig. 34. For a detailled description see Fig. 44 and 45)

not located in the periphery, but originating in the center of the FU, a significantly higher resolution would have to be reached. Two explanations can be advanced to explain this effect. It is logical that a well pronounced alpha helix outside the main density is easier to "see" than one buried within other structures. A second point is also significant. Since the FUs have changed their spatial orientation, an overlay of identical FUs in different conformations occurs and accordingly the exact identification, where such shifts happen is difficult. The last point, the binding of allosteric effectors, cannot be visualized using a method which has not yet reached atomic resolution. The interpretation of a molecular model and cryo-TEM map can, however, provide more information than the density map alone, making the conformational changes interpretable.

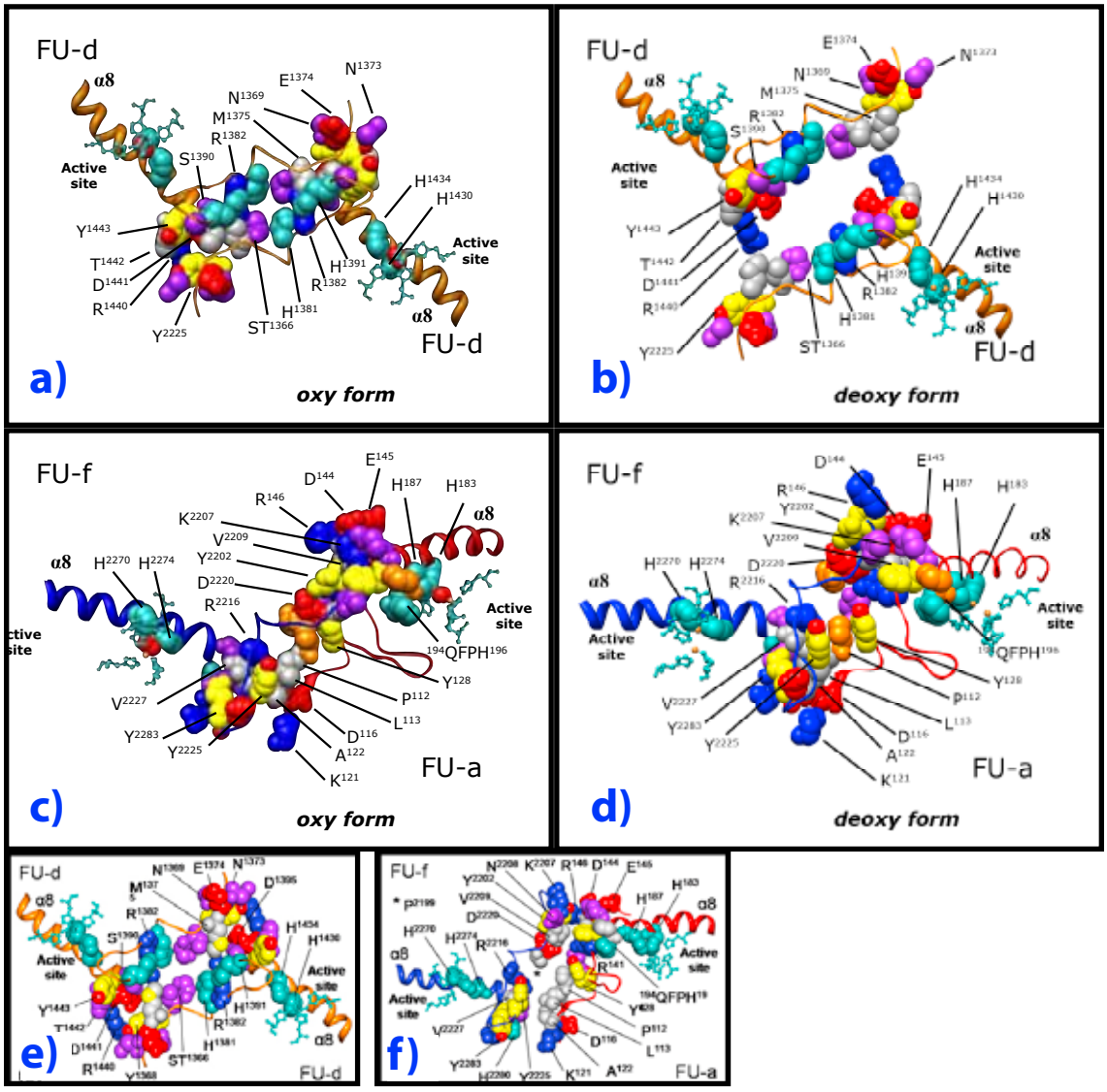


Fig. 37: Major groove interfaces of NpH oxy/deoxy. Position and orientation of interfaces are shown in comparison to the published model (Gatsogiannis et al. 2007). (Color code see Fig. 34. For a detailled description see Fig. 47 and 48)

Any cooperative behavior of the hemocyanin molecule has to be transferred between the FUs, over interfaces. Thus, already a careful description of the molecular model with detained attention to these interfaces has been performed (Gatsogiannis et al. 2007), see above (see Fig. 34-39). Since this description is based on a dataset gained by a preparation containing mixed oxygenation conditions, this data cannot be used for any conformation-specific interpretation. Thus, the data form the first publication has been modified in order to describe the allosteric interactions, on basis of the new oxy and deoxy NpH model. As already shown in the previous publications on NpH (Gatsogiannis et al. 2007) and KLH1 (Gatsogiannis and Markl 2008) the predominant secondary structure elements (alpha helix α15 and the six stranded anti-parallel β-sandwich)

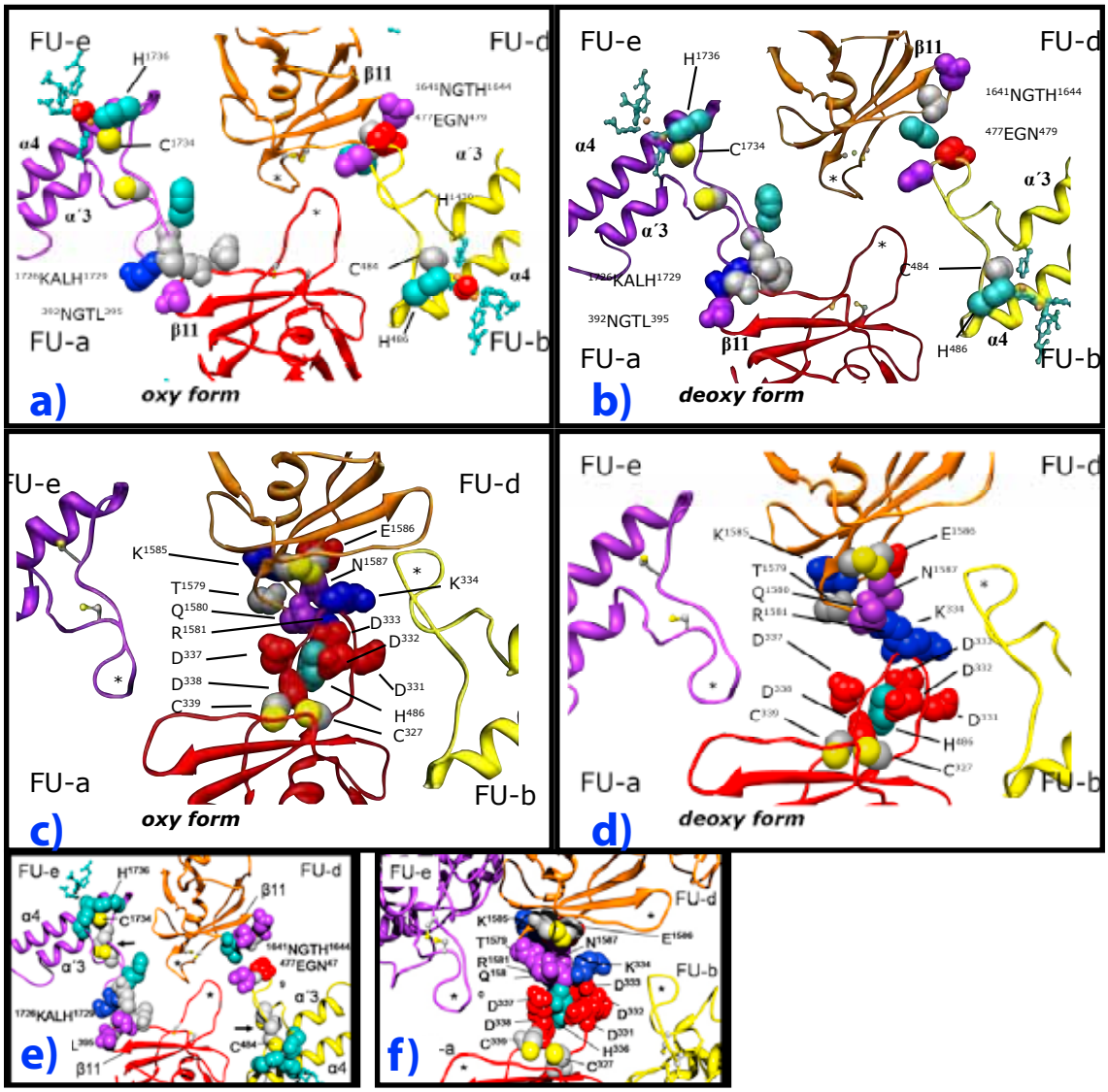


Fig. 38: The anchor interface of oxy and deoxygenated NpH. Position and orientation of interfaces are shown in comparison to the published model (Gatsogiannis et al. 2007). (Color code see Fig. 34)

are easy detectable. This leads to an unambiguous rigid body fitting of the homology-modeled FUs into the two datasets. Since the goal was a straightforward comparison of the two conformations, neither loop refinement nor flexible fitting have been performed. Furthermore, no rotamers of the amino acids have been altered. Such refinement steps bear a deep risk of over interpretation which should be strictly avoided. Special care was taken relating to the densities between two adjacent FUs within the molecular models. Thus, in some cases new residues were added to the existing list (see Fig. 33). As already described many interfaces offer positions for possible salt bridge interaction of amino acids. These are difficult to firmly establish on two partner residues of opposite charge. However, the conspicuous distribution of charged residues at inter-FU contacts, together with tube-like densities at exactly these locations are good indicators for salt bridges.

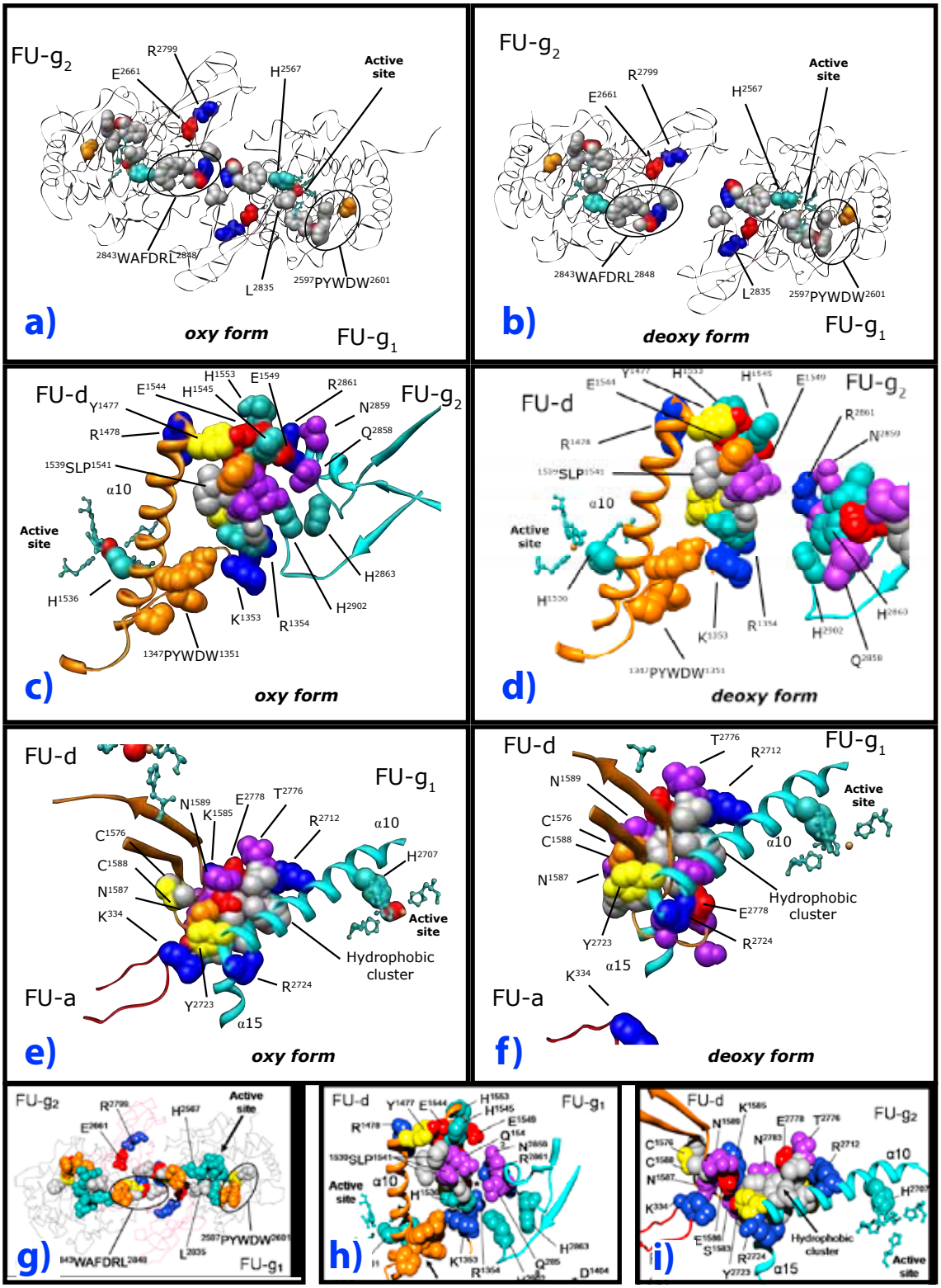


Fig. 39: Collar-interfaces. Morphological unit interface g↔g connects two arc FUs g inside the ring structure (a+b). The remainig interfaces d↔g2 (c+d) and d↔g1 (e+f) reflect the connections of arc and wall. Position and orientation of interfaces are shown in comparison to the published model (Gatsogiannis et al. 2007). (Color code see Fig. 34)

Such a density could evolve, when the rotamers of an amino acid are oriented in such way that they point at each other in the majority of particles observed. This serves as a further argument for salt bridging.

The 12 types of wall interfaces can be distributed into four main interface groups. These are three different connections along the horizontal axis as well as one connecting the peripheral tier with the middle tier. As a common feature, these interfaces interconnect the long α-helices, from the adjacent FUs, carrying active site histidines at the distal ends. This arrangement is considered to play an important role in allosteric signal transfer.

The morphological unit interface (a↔b, d↔e, c↔f, g↔g) (see Fig. 34 and 40-42) has the same appearance as the organization of the two FUs in the crystal structure of *Octopus* FU-g (Cuff et al. 1998). In this publication the basic features of this interface were presented and have been subsequently adapted to NpH (Gatsogiannis et al. 2007). In these interfaces the FUs are associated at their β-sandwich domains in an anti-parallel manner. Referring to the interconnecting densities, two main bridges can be identified. One is located at the top, where the linker is passing from one FU to the other (no linker in case of c↔f) and another broad contact zone below. The upper connection is a very good candidate for salt bridging, since it contains charged residues at relative close distance. A tube like density bridges the gap between the FUs at this position. This is a strong indication for a salt bridge formation, because in this case the rotamers of the residues taking part are clearly directed towards each other. This leads to a stable bridge in the density produced by cryo-TEM. While the organization of this bridge is essentially the same among all interfaces belonging to this group, the broad contact zone varies. Here, the WX(F/Y)DRL motif, which each β-sandwich domain contains, might also play an important role in signal transfer. These highly conserved sequences oppose each other at the center of the morphological unit interface. Whereas in a↔b we find multiple charged amino acids together with two histidines, the other two cases d↔e and c↔f have 3 and 4 histidines respectively, but still there are pairs of charged residues around this density. The charged cluster provides many possibilities for salt bridge interaction, whereas the histidine-rich motifs could further help to mediate allosteric interaction. In a previous study on the highly cooperative Limulus hemocyanin, clusters of histidines have been found to be associated with transmission of allosteric interaction between the eight hexamers (Martin et al. 2007); a similar mechanism could transfer allosterism in NpH. The listed features posses the capability to transfer signals. Their structural differences suggest that they might play distinct roles in allosteric interaction, a feature supported by the fact that in *Octopus* hemocyanin active site functional heterogeneity

has been observed (Miller 1984; Connelli et al. 1989; Gatsogiannis et al. 2007).

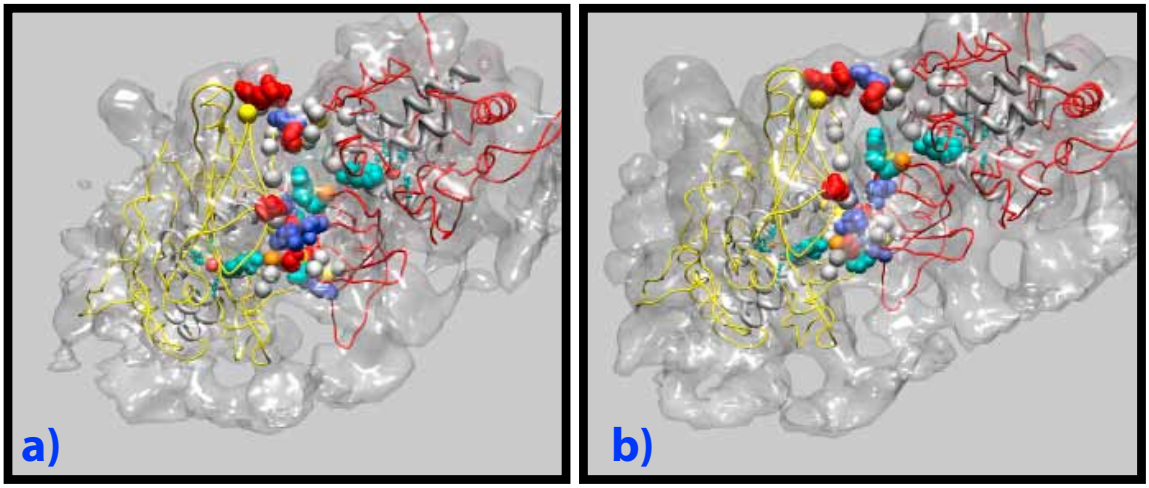
Interfaces d↔d and a↔f in NpH (see Fig. 37 and 47+48) also have highly conserved features. At first glance it is clear that the difference between both functional states is the largest here. This is due to the significant movement of the FUs seen when comparing both conformational states. FUs f and d in particular show the largest changes in spatial orientation. This change is most obvious with respect to the alteration of the distance between the loops that connect alpha helices α5 und α6. The principle architecture is a central density at the core domains of the adjacent FUs. Alpha helix α8, carrying an active site histidine, ends very close to the interface. Thus this alpha helix might transfer a signal, as discussed already. The long loop between alpha helices α5 und α6 offers possibilities for salt bridging and hydrophobic interaction, and seems to play a significant role as the cause of reorientation. As can be seen in Fig xy and Figxy the gap between the neighboring FUs is significant in the deoxygenated state (in both cases), whereas in the oxygenated conformation it changed to a small slit. This indicates that the oxy conformation is more stable in this case. In the center of d↔d, many histidines are again located.

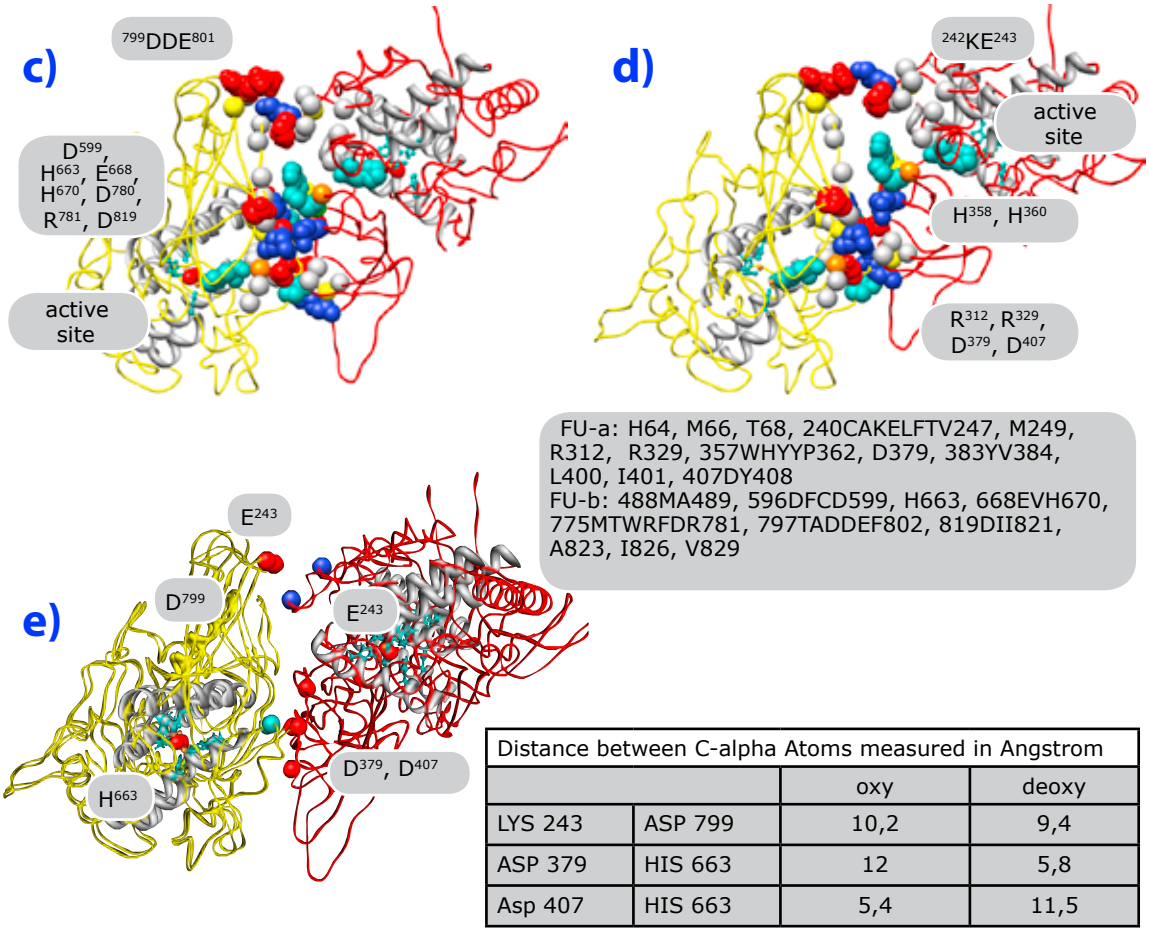
The so called, minor groove interfaces b↔c and e↔e (see Fig. 36 and 44+45) have two symmetrical contact sites at their distal ends. Amino acids of the, core- and beta sheet domain connecting, α15 helix contribute to the charged cluster at each side. In the center of this interface is a density-free window. The clusters provide several possibilities for chemical bonding by salt bridges and hydrophobic interaction. The invariant motive PYWDW is situated in the middle of both bridges. This motive is associated with alpha helix α10, connected to the active site histidine and is believed to influence oxygen-binding behavior (Cuff et al. 1998; Miller et al. 1998). The overlay of both NpH conformational states shows that alpha helix α15 is shifted along its axis between the two conformations. This helix, acting as hinge, can therefore influence the contacts between both domains supporting allosteric signaling across the interface.

Interfaces f↔e and d↔b (see Fig. 35 and 46+43) connect the peripherical tier with the middle tier. A comparison between the organization in oxy- and deoxy- NpH form directly shows that both interfaces are narrow in deoxy state and broader in the oxy state. In addition, there is similarity as both interfaces are composed of many charged residues, offering many possibilities for salt bridge interaction. The α4 helixes point directly into these clusters and can therefore transfer signals between the adjacent active sites.

In OdH it was found that seven active sites encompass one allosteric subunit (Miller 1984; Zolla et al. 1985; Miller et al. 1988; Connelli et al. 1989). Since *Nautilus* is closely related to *Octopus,* the situation is likely to be identical. Therefore, 10 allosteric subunits, containing one of each of eight FUs, should form the decameric molecule. As already described, the morphological unit interfaces offer ideal preconditions for inter FU signal transfer. This makes it very likely that one of each copy is included within the allosteric subunit. The interfaces connecting the tiers (b↔d, e↔f) (see Fig. 35 and 43+46) are also rather promising in this aspect. Therefore, the findings of this work support the suggestion already made (Gatsogiannis et al. 2007), that the allosteric subunit encompasses the densely packed series of six wall FUs that lie between the major and minor groove. The cooperativity of a molluscan hemocyanin is lost upon dissociation into its subunits, while the oxygen affinity increases (Bonaventura et al. 1981). This means that cooperative oxygen binding depends on a fully assembled oligomeric protein. Taking this into account it is very likely that an allosteric subunit has in some way to *work* against a counter part in order to exhibit the cooperative oxygen binding behavior. This could be established by the remaining interfaces. They, although not belonging to the allosteric subunit, also have strong interaction capabilities with their adjacent FUs. It is very likely that the neighboring allosteric subunit is working as a counter surface. Since the assembled protein is essentially a ring, this support would encompass the whole molecule, producing self-stabilization.

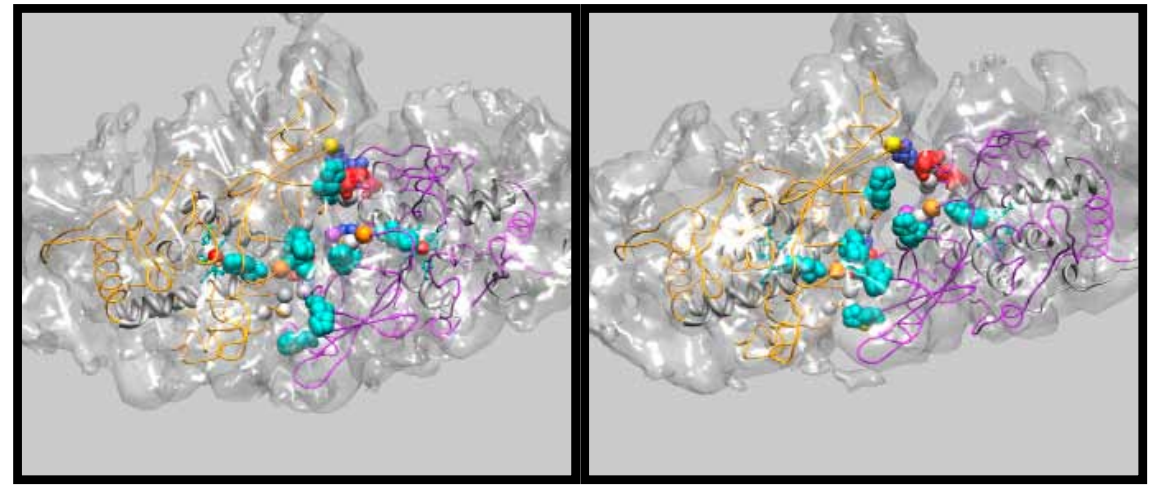
Cooperative proteins have to have counter parts, as a general feature. This is a prerequisite for the classical MWC-model (Monod et al. 1965) of cooperative oxygen binding. Though this basic model needs to be modified to account for all hemocyanins (Decker et al. 2007), it efficiently describes the O_2 binding behavior of NpH (Miller 1984). Another requirement for the MWC model is that at least two states, which maintain the molecular symmetry, exist. These states should differ in distribution and/or energy within the inter-subunit bonds. This would apply conformational constraints upon the architecture of one allosteric subunit. All these assumption can easily be transferred by the features found and described here. As already mentioned, salt bridge interaction between adjacent FUs is highly likely, according to the interface architecture and the contributing amino acids. Following Coulombs law, the energy of such a bond is reduced reciprocally by the distance of the partners, by the power of two. Thus, a small variation causes a significant change in energy. The fact that changes in ion concentration have strong effects on oxygen binding behavior also support the assumption that salt bridges are involved in this process. Derived from the distance measurements obtained, it is apparent that one bridge can only stabilize one conformation. Thus by rearranging the spatial orientation of the FUs, the distribution of energy within the bonds of one allosteric

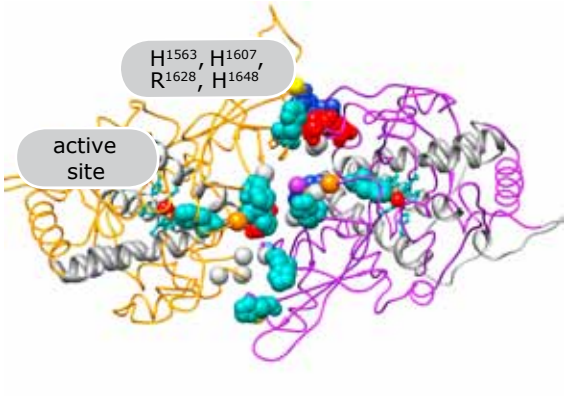


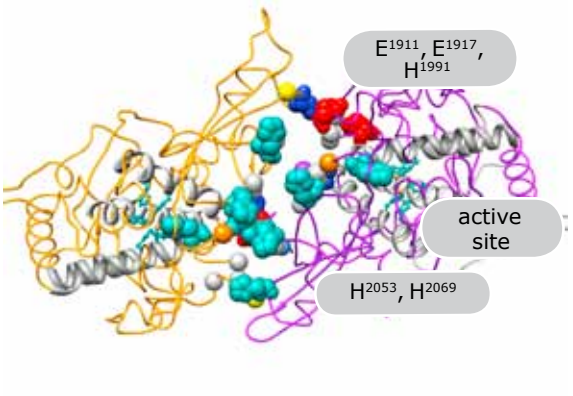


Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
LYS 243	ASP 799	10,2	9,4
ASP 379	HIS 663	12	5,8
Asp 407	HIS 663	5,4	11,5

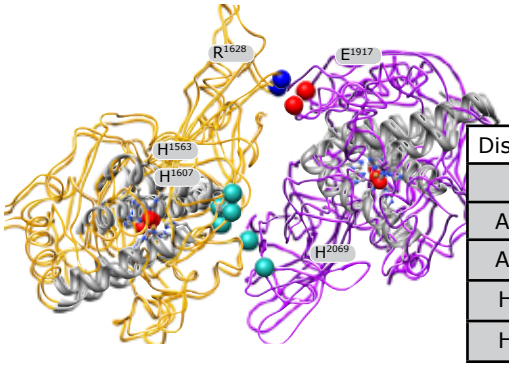
Fig. 40: Morphological unit interface a↔b in two conformations (FU-a red, FU-b yellow). The extracted densites with their respective homology models (a+b) reflect the high quality of the fitting. Note the two densities at the top and in the center. The upper provides possibilities for salt bridges and is enforced by the linker peptide between FU-a and FU-b. The center is composed of multiple charged residues and histidines, which allow signal transfer between the two active sites over the α helix 8 coupling. The overlay of both FUs reflect the relative change of spatial orientation (e). While FU-b is stable the variation is significant in FU-a. (Color code see Fig. 34)





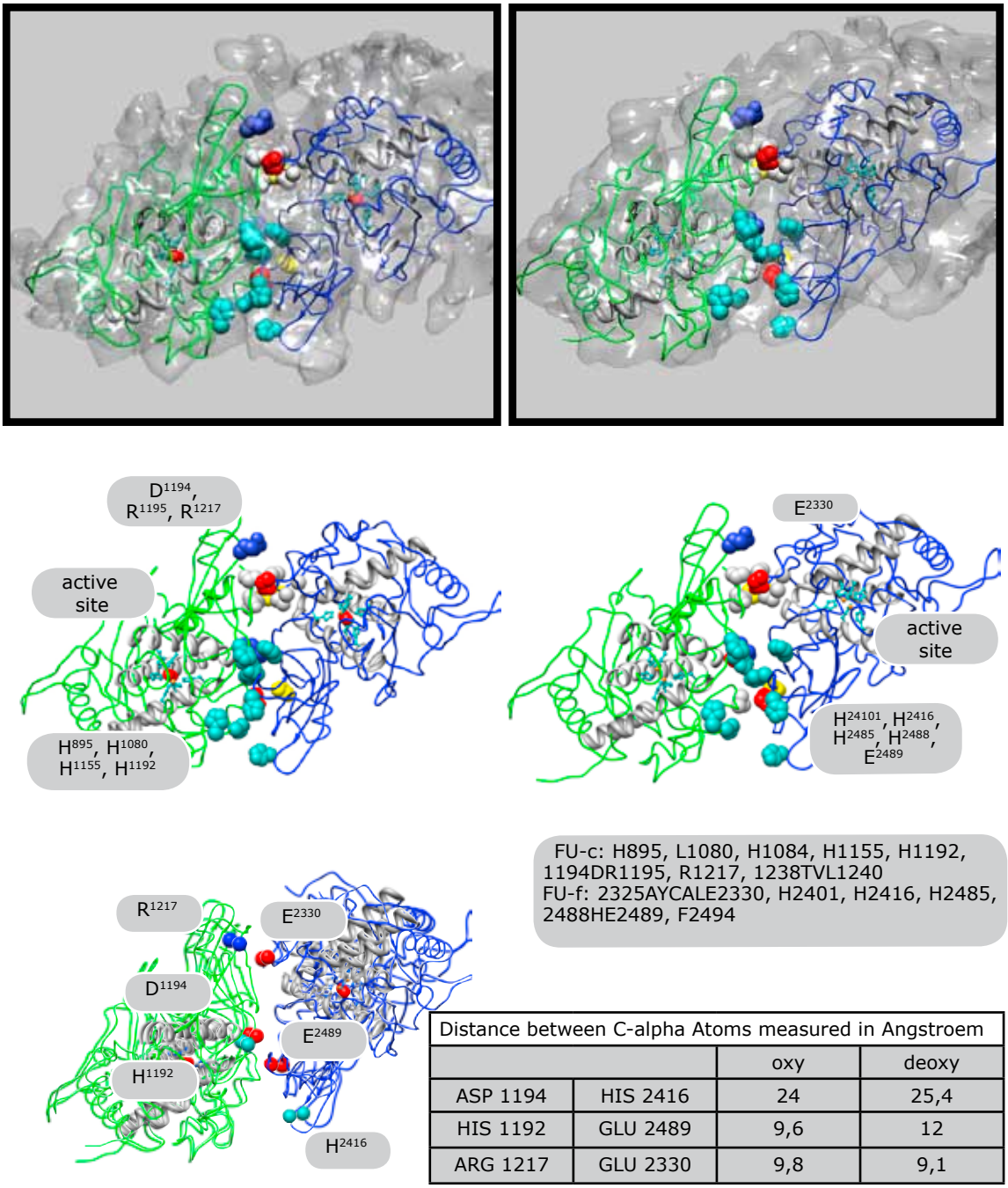


FU-d: 1488AL1489, M1492, H1563, H1607, R1628, F1631, H1648, I1655
FU-e : E1911, 1914CALE1917, H1991, 2052FH2053, H2069, L2077, P2080



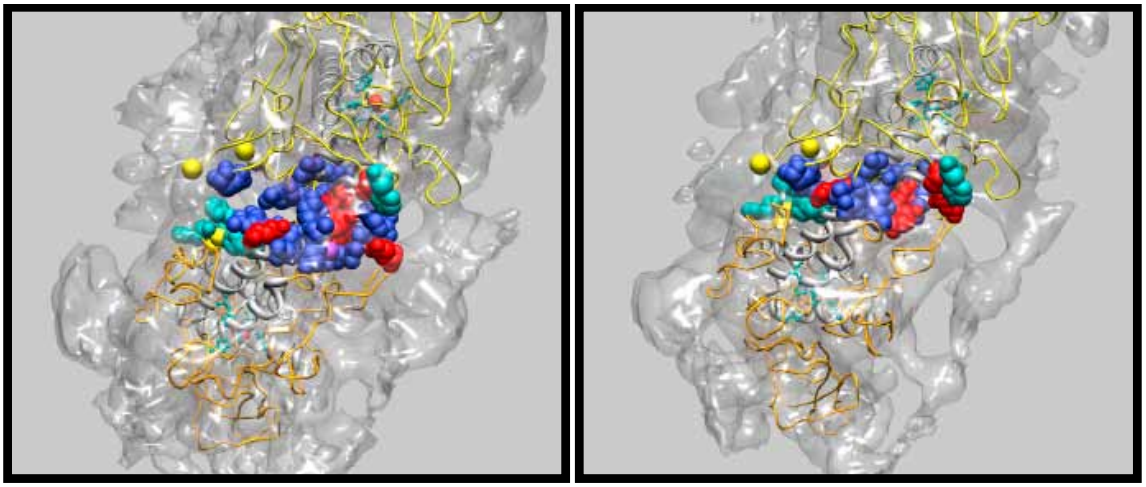
Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
ARG 1628	GLU 1917	10,1	6,7
ARG 1628	GLU 1911	10,6	15,7
HIS 1563	HIS 2069	15	8,8
HIS 1607	HIS 2069	12,6	8,8

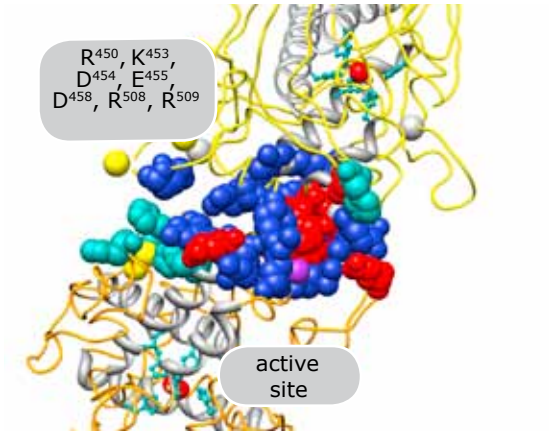
Fig. 41: The central tier morphological unit interface d↔e (FU-d orange, FU-e purple) has a "tube like" density at the top. Within this density are charged residues indicating salt bridge formation. In the center three histidines (H1563, H1607 and H1991), together with three charged residues (D1609, R1610 and R2031), might transfer allosteric signals. (Color code see Fig. 34)

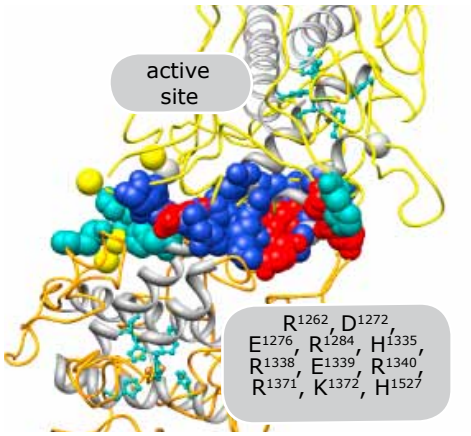


Distance between C-alpha Atoms measured in Angstroem			
		oxy	deoxy
ASP 1194	HIS 2416	24	25,4
HIS 1192	GLU 2489	9,6	12
ARG 1217	GLU 2330	9,8	9,1

Fig. 42: Morphological unit interface c↔f (FU-f blue, FU-c green). At the upper part are four charged residues (D1194, R1195, R1217 and E2330) within a hydrophobic cluster. In the center of this interface several histidines might play an important role in signal transfer. Additionally a pair of two charged residues (D1194 and R1195, FU-c and D2440 and R2441, FU-f) are in close contact with an active site α helix. In this case the reorientation of the FUs is not very strong (see table Asp 1194/His 2419). However, signals could be transferred using the contacts as coupling and the helices as hinge between the two active sites. (Color code see Fig. 34)



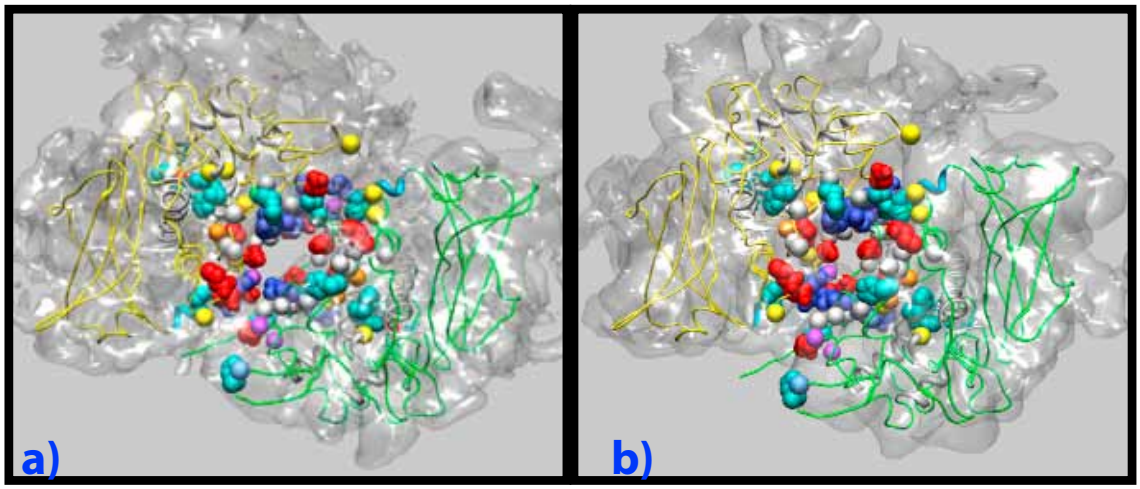


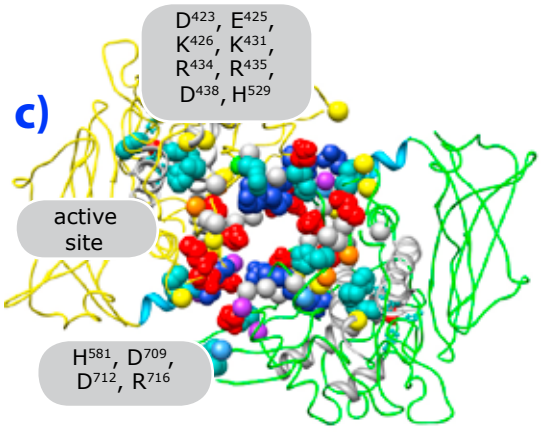


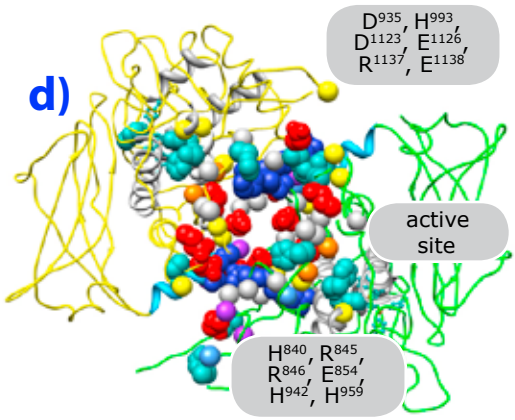
FU-b: R450, 453KDESID458, R508, R509
FU-d: R1262, D1272, E1276, N1280, R1284, H1335,
1338RER1340, R1371, K1372, H1527

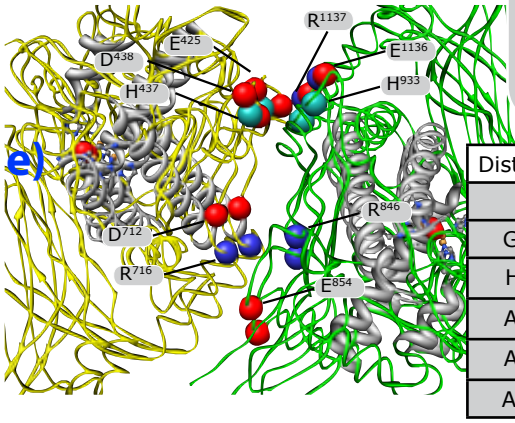
Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
ARG 508	GLU 1274	8,5	5,7
ARG 450	GLU 1339	8,5	4,1
GLU 455	ARG 1279	8,5	5,9

Fig. 43: Horizontal interface b↔d (FU-b yellow, FU-d orange) between the peripherical and central tier. The broad contact zone offers many opportunies for salt bridges. Especially the central arginine-rich cluster seems to be important. The distance between the FUs is significantly reduced in the deoxy-conformation. Thus, the gap, in the oxy-form density, is closed. Also note the histidine cluster at the left side of the interface (FU-d). The two α4 helices connect the active sites and are involved in the interface architecture. (Color code see Fig. 34)





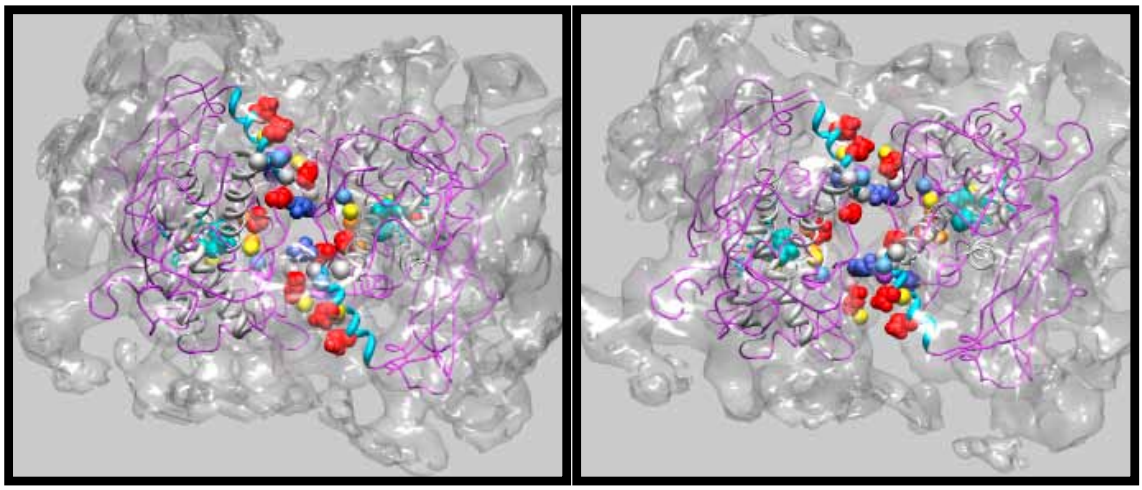


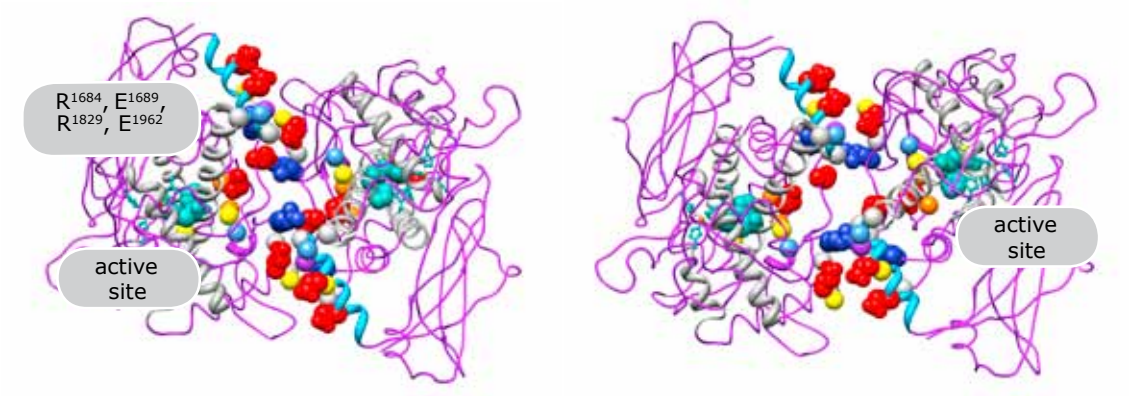


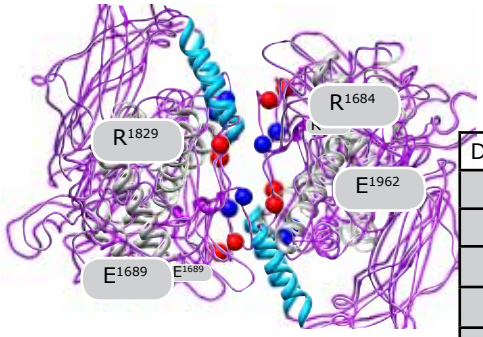
FU-b: D423, E425, K426 K431, 434RRL-HD438, 524MI525, L527, 529HL530, F578, H581, 709DTLDF713, R716
FU-c: H840, 844TRR846, I848, 851LT852, E854, D935, 939AITH942, H956, F991, 993HF994, Y1122, D1123, 1125LE1126, L1135, 1137RE1138

Distance between C-alpha Atoms measured in Angstroem			
		oxy	deoxy
GLU 425	ARG 1137	12,7	6,7
HIS 437	GLU 1138	15,6	12,2
ASP 438	HIS 993	10,1	8,2
ASP 712	ARG 846	14,3	14
ARG 716	GLU 854	13,5	17,5

Fig. 44: The minor groove interface b↔c (FU-b yellow, FU-c green) has several opportunites for chemical bonding. Note the contact-sites surround a cavity (a+b). This cavity is smaller in the deoxy-conformation (b). Note the possible active site interactions over the α helices 15 and 10 involved in this interface formation. While α helix 10 is an active site helix, α helix 15 connects the core and β sheet domain. (Color code see Fig. 34)



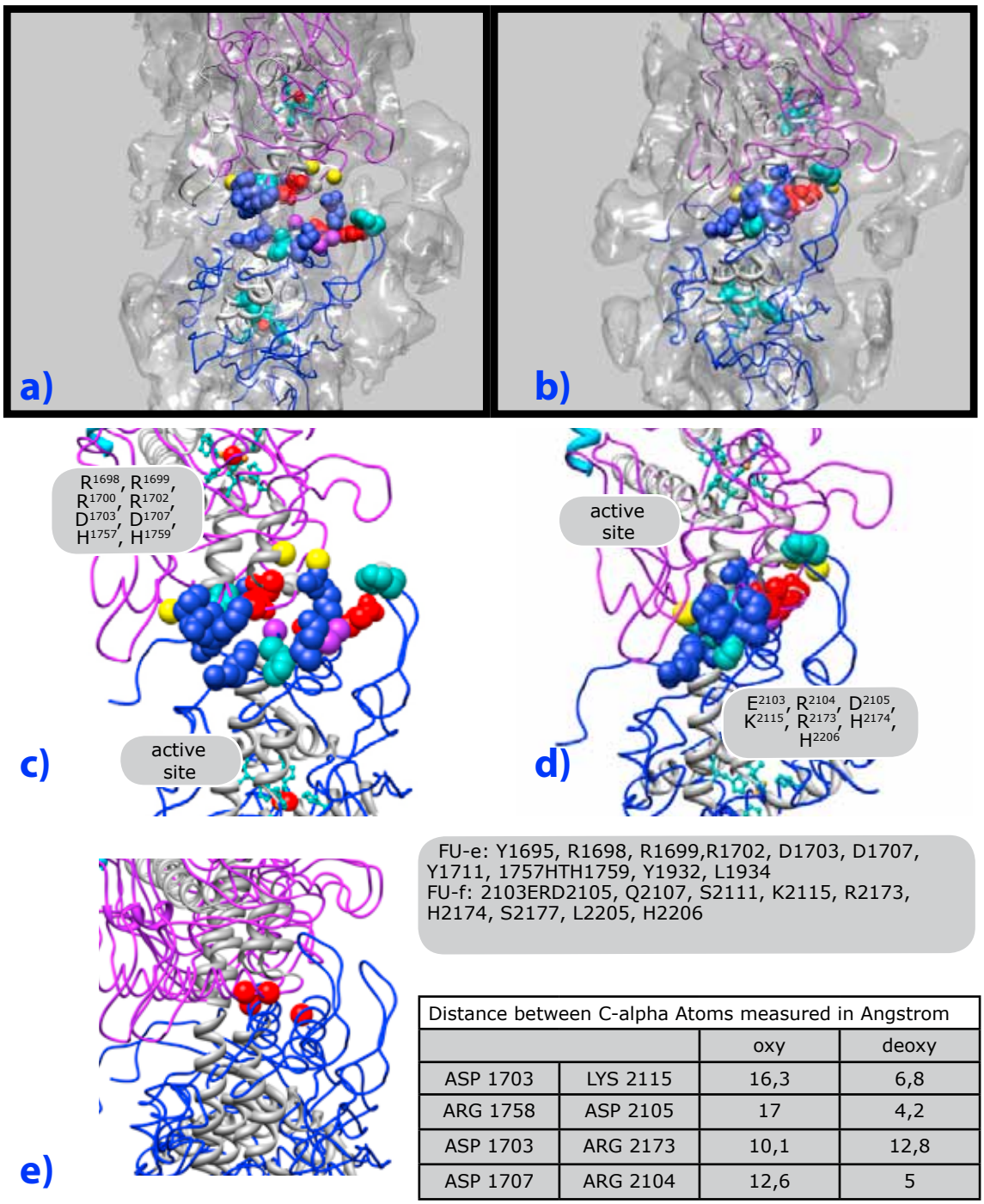




FU-e: 1684RL1685, 1687IYE1689, R1829, E1962, M1966, Y1974

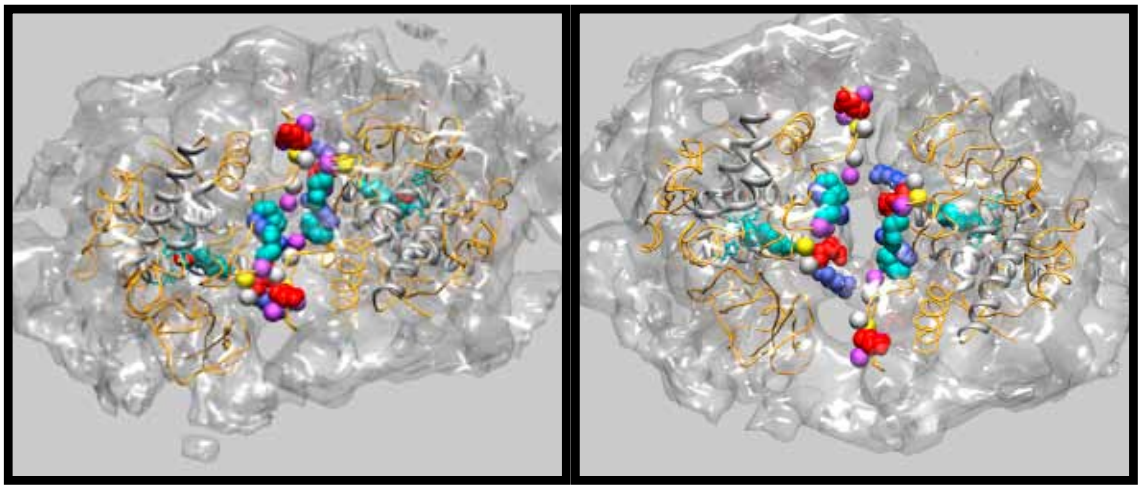
Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
GLU 1689	ARG 1829	11,5	15,9
ARG 1684	GLU 1962	8,8	14,4
ARG 1829	GLU 1689	8,8	15,7
GLU 1962	ARG 1684	13,4	14,9

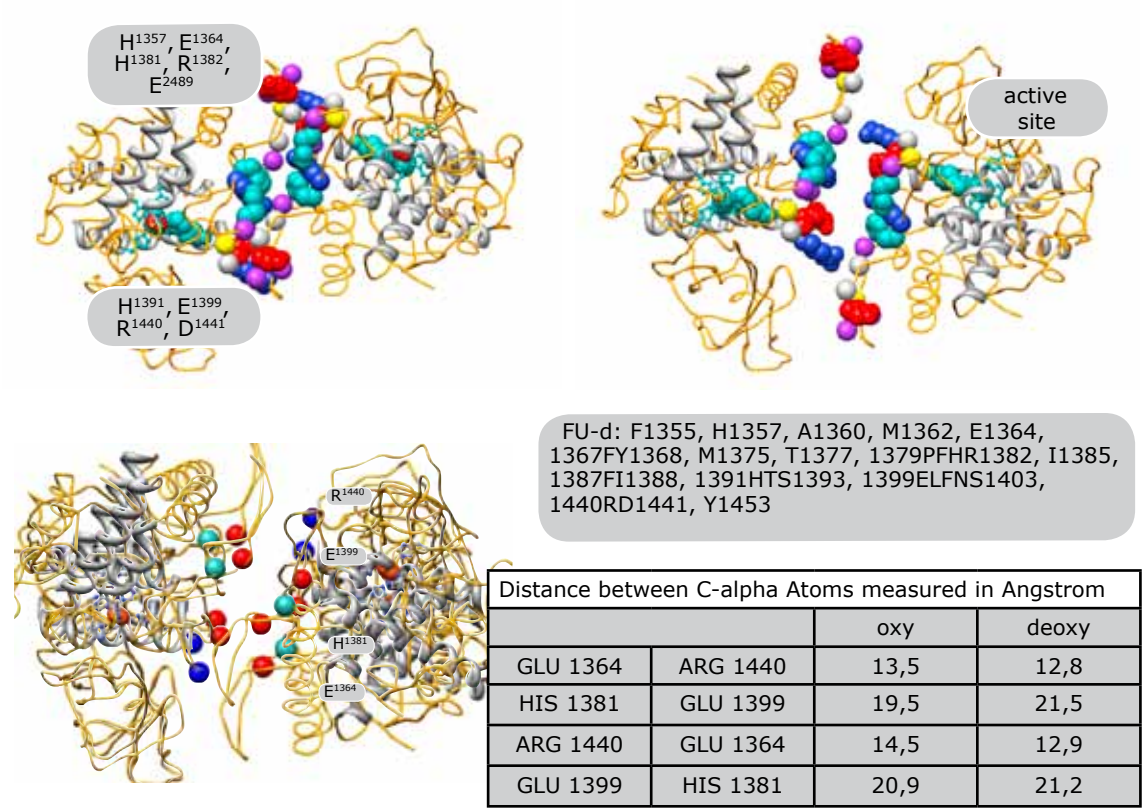
Fig. 45: The minor groove interface e↔e, in the central tier, has a cavity in the center . The contact zones are postitioned at the ends of α helix 15. This helix connects the beta sheet part of the FU with the core domain. Shifts of this helix might be transferred to the core domain over hinge movement. (Color code see Fig. 34)



Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
ASP 1703	LYS 2115	16,3	6,8
ARG 1758	ASP 2105	17	4,2
ASP 1703	ARG 2173	10,1	12,8
ASP 1707	ARG 2104	12,6	5

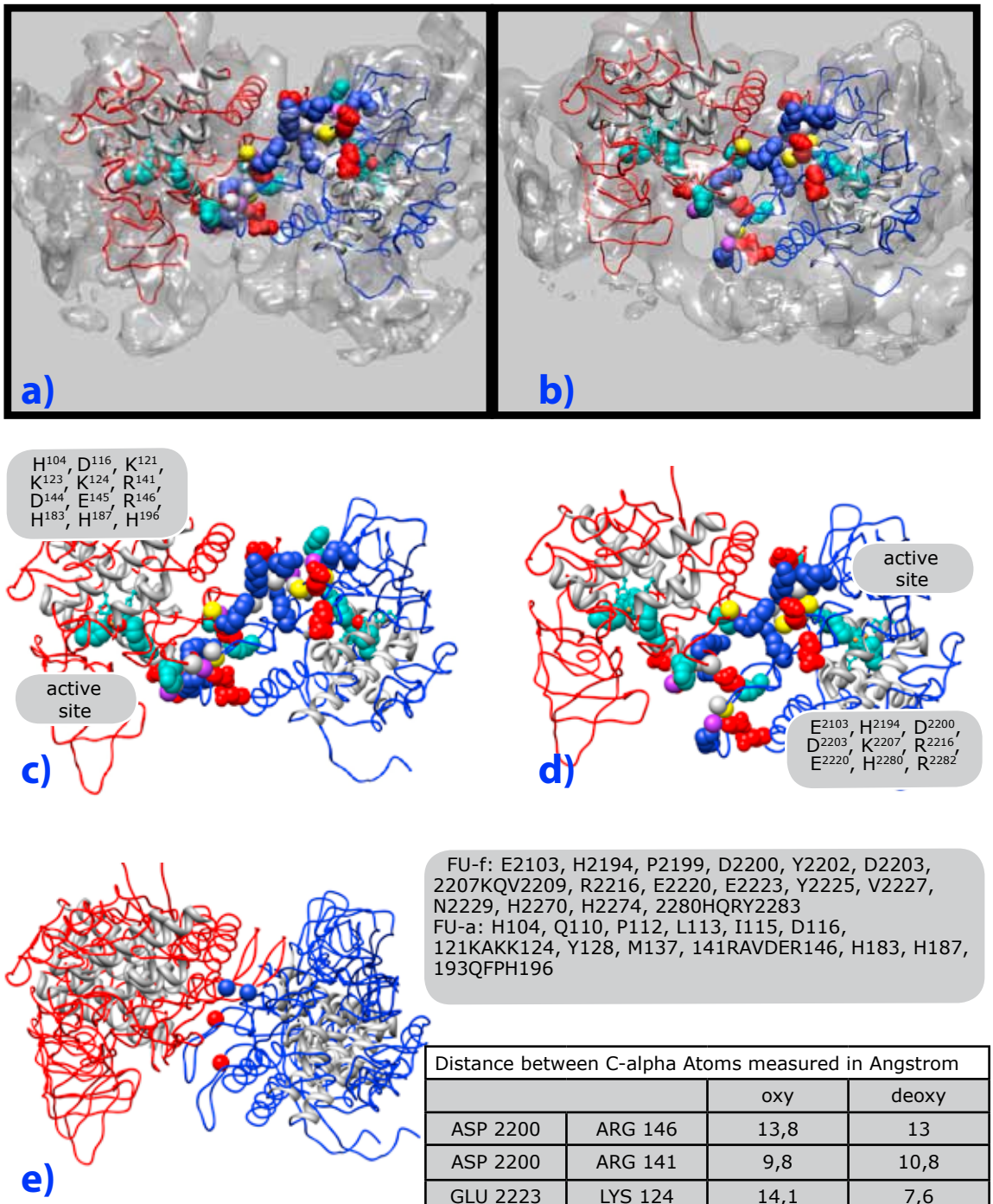
Fig. 46: The horizontal interface f↔e (FU-f blue, FU-b purple) shows significant changes between both conformations (a+b+e). The arginine-rich interface is pushed together in the deoxy conformation (d). This is mainly due to the significant movement of FU-f (e). Note the distance variations between both conformations. (Color code see Fig. 34)





Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
GLU 1364	ARG 1440	13,5	12,8
HIS 1381	GLU 1399	19,5	21,5
ARG 1440	GLU 1364	14,5	12,9
GLU 1399	HIS 1381	20,9	21,2

Fig. 47: Central tier interface d↔d shows strong movement between the two conformations. As shown in the table the distance variation is equally distributed. In the deoxy-form a significant cavity can be found in the center of this interface, which is closed in the oxy-state. The broad contact provides several spots for chemical bonding. The acitve site helix α 8 could transfer a signal of oxygenation to the neighboring FU via a hinge movement of the broad contact zone. (Color code see Fig. 34)



Distance between C-alpha Atoms measured in Angstrom			
		oxy	deoxy
ASP 2200	ARG 146	13,8	13
ASP 2200	ARG 141	9,8	10,8
GLU 2223	LYS 124	14,1	7,6
ARG 2282	ASP 116	13,3	6,8

Fig. 48: Molecular structure of the a↔f interface at the major groove (FU-a red, FU-f blue). This interface provides several sites for chemical bonding. The density distribution varies due to the spatial orientation of the FUs in both conformations. While in the deoxy conformation (b) a cavity is present in the center of the interface, this dissappeares in the oxy form (a). Note the significant movement of FU-f ((e) and distance measurements). (Color code see Fig. 34)

subunit is altered, as demanded by the MWC model. The fact that an isolated subunit has a higher oxygen affinity than the assembled molecule is, for similar reasons, easily explainable. In a fully assembled decamer, neighboring subunits also apply forces *via* the interfaces. Again spatial rearrangement of FUs causes a redistribution of energy. This has two effects. The actual conformation of the molecule is stabilized which is the reason for the high oxygen affinity of an isolated allosteric subunit. The second effect is that information on the oxygenation state of one subunit is transferred to the neighboring FUs, resulting in cooperative oxygen binding behavior. This was already mentioned in the previous *Nautilus* publication (Gatsogiannis et al. 2007), where a two-step organization on the basis of a Nested MWC model was presented. This Nested MWC model is a hierarchical extension of the classical MWC model. The horizontal signal transfer across the major and minor decamer groove might force a subunit to reorganize into the other conformation, if the neighboring subunit has already changed. Again, this would be transferred through changes of salt bridge bonding, caused by spatial orientation due to FU rearrangement. This would then account for the Nested MWC explanation, which allows for two-step cooperative oxygen binding behavior. Assigning this two-step effect to the NpH molecular model introduced here is relatively straightforward. After the cooperative oxygen binding has occurred to one allosteric subunit, the changes of interface architecture apply forces to the neighboring subunits, although the subunit itself does not change its architecture significantly. If the partial O_2 pressure is high enough, more than one subunit will bind oxygen. The inter-subunit pressure will rise, until enough subunits have bound O_2. At this point the whole molecule will change its conformation into the oxy-state, although not all of the subunit will be fully oxygenated. In this model, both functional states are stable, but can easily be reversibly converted into each other. The principle of this mechanism is simple, but still satisfies all the demands required of the MWC model. Nevertheless, hierarchical oxygen binding as described in the Nested MWC model has not yet been found experimentally. It is likely that it would be worthwhile to carefully evaluate the data with respect to this question. Accordingly, the model remains speculative.

However, the two NpH conformations clearly show structural changes due to FU reorientation. It is highly likely that these changes are caused by rearrangement of inter-FU salt bridges. Another fact supporting the model is that it fulfils all requirements of the classical MWC model. A final supporting argument is that the oxygenation-dependent allosteric conformational change in human hemoglobin, as the most intensively analyzed protein, is also mediated by salt bridges.

Conclusion and Outlook

My thesis focused on two topics. First to reveal the quaternary structure of SoH and to deduce a *pseudo* atomic model from it. This in turn would show how the additional FU-d' is included into the molecular assembly. The second part was the development of a method capable to reliably show NpH in two different conformations (ie oxy- and deoxygenated state). The results should then be interpreted by means of allosteric interaction and conformational changes due to cooperative oxygen binding behavior. As a spin-off, this study helped to reveal the quaternary structure of NpH and to built the first *pseudo* atomic model of a molluscan hemocyanin (NpH, published 2007).

The work on SoH showed that the wall architecture is identical compared with the other molluscan hemocyanin types. Here the first six FUs (FU-a to FU-f) form the cylindrical wall, while the additional seventh and, if existing, eighth FU (FU-g, FU-h) built an internal arc structure. The eight FU in SoH is not appended at the end of the polypeptide chain, as it would be in other eight FU-type decamers, but is included in the middle of it. This is due to a duplication event of FU-d (FU-d'). Prior to this work it was unknown whether the wall is formed by a linear progression of FUs (ie FUs a, b, c, d, d', e would form the wall FUs f and g the arc), or if the additional FU-d' is included in arc architecture. Following the hybrid approach of rigid body fitting and putative linker distance measurements, it was revealed that FU-d' is included in the arc, proving that the wall architecture is fundamental to all molluscan hemocyanins. The interpretation of the molecular model of SoH again showed high similarities with NpH by means of interface architecture.

The oxy/deoxy experiment on NpH proved that it is possible to fix different conformations, caused by modulation of atmospheric conditions, in vitreous ice. Prior to this study this fact was intensively discussed and doubted. It was stated that, due to the vacuum in the TEM, the hemocyanins loose all their oxygen, so that only deoxygenated hemocyanins may be observed by electron microscopy. Since no direct measurement of the oxygenation state of the specimen on the frozen grid can be obtained, it was very challenging to prove that this approach is routinely applicable. Performing a cross-over experiment it was shown that the two different datasets reflect different conformations of the hemocyanin. This gave evidence, that the differences observed are coupled with the differences in specimen preparation (ie oxy/deoxy). The analysis of the data gained by this experiment showed significant changes in spatial FU orientation between both con-

formations. This in turn results in rearrangement and reorientation of inter FU interfaces. Densities within the gap of two neighboring FUs almost always coincided with cluster of charged amino acids. These densities could be explained by salt bridges. Two oppositely charged residues should point at each other. This in turn would result in a tube-like density within the gap of two FUs. Since the distance between these amino acids differs in both states (caused by spatial reorientation of the FUs) such a bridge can only stabilize one conformation and not the other. A balanced distribution of such bridges around the molecule could stabilize at least two conformations of the hemocyanin. The fact that salt bridges play an important role in cooperative oxygen binding is supported by the finding that certain ions can significantly change the oxygen affinity of molluscan hemocyanins. Based on the classical MWC-model of cooperativity a model of how cooperativity might in NpH work was supposed.

This work paved the way for intensive studies on oxygen dependent conformational changes of hemocyanins. If it is possible to reach a resolution where all secondary structure elements are clearly visible the studies of how cooperative interaction is happening can be expanded. This should give even more insight into the allosterism of this molecular machine. Furthermore it should be observed how the bohr effect (pH depended change of oxygen affinity) affects the conformation of the molecule. Additionally this method should also be applied to arthropodan hemocyanins which exhibit the strongest cooperativity of all oxygen binding proteins yet observed. Finally the significant changes observed in NpH imply that it is obligate to control the atmospheric conditions if a protein affected by different gaseous environments is investigated. Otherwise high resolutions can not be obtained due to averaging of different conformations.

Summary

The hemocyanins of the cephalopods *Nautilus pompilius* and *Sepia officinalis* facilitate the oxygen transport between the gills and the consuming tissues of these marine predators. They appear as cylindrical decamers with internal arc structure. While one subunit (ie one polypeptide chain) inhabits seven paralogous functional units (FUs) in NpH (termed FU-a to FU-g), a gene duplication event of FU-d led to eight FUs in SoH (termed FU-a, b, c, d, d', e, f, g). In all molluscan hemocyanins six of these FUs form the external wall, while the remaining constitute an internal arc structure.

The here presented 3D-reconstruction of *Sepia officinalis* hemocyanin (SoH) enlarges the understanding of structural molluscan hemocyanin assembly. The 8.8Å map (FSC=0.5) allowed to obtain a molecular model with *pseudo* atomic resolution by homology modeling and rigid body fitting. Furthermore the two reconstructions of *Nautilus pompilius* hemocyanin (NpH) under different states of oxygenation gave insight into conformational changes caused by oxygen binding. With 10 and 8.1Å (FSC=0.5) they provide the currently highest resolution of a molluscan hemocyanin decamer and their analysis allowed to propose a model for allosteric signaling occurring during oxygenation.

The molecular model of SoH revealed the topology of the eight FUs, showing that FU-d' is integrated in the arc and that the principal wall architecture is common to all different types of molluscan hemocyanins. Furthermore the model shows that the molecular interface architecture, of both isoforms of SoH, is almost indistinguishable. The analysis of the molecular model revealed high similarities compared with NpH. Additionally, the molecular model allowed to withdraw a lot of information concerning signal transfer and function within this molecule.

The pioneering structure-to-function approach applied to NpH oxy- and deoxy conformation shows that the experimental setup of preparing the specimen under two different atmospheric conditions is successful and can now be routinely applied on other proteins. The high quality of the data gained allowed the interpretation of the two conformations on basis of their molecular models. The comparison of the 3D-densities showed that the FUs slightly change their respective spatial orientation upon oxygenation. Thus, affecting interface architecture and distance especially between opposing amino acids. These findings led to a model, of how allosterism can be transferred between adjacent FUs and subunits, based mainly on salt bridge rearrangement, due to changes of spatial orientation.

Zusammenfassung

Die Hämocyanine der Cephalopoden *Nautilus pompilius* und *Sepia officinalis* sorgen für den Sauerstofftransport zwischen den Kiemen und den Geweben. Sie bestehen aus einem zylindrischen Dekamer mit interner Kragenstruktur. Während eine Untereinheit (also eine Polypeptidkette) bei NpH aus sieben paralogen funktionellen Domänen (FU-a bis FU-g) besteht, führte ein Genduplikationsereignis der FU-d zu acht FUs in SoH (a, b, c, d, d´, e, f, g). In allen Mollusken Hämocyaninen bilden sechs dieser FUs den äußeren Ring und die restlichen die interne Kragenstruktur.

In dieser Arbeit wurde ein dreidimensionales Modell des Hämocyanins von *Sepia officinalis* (SoH) erstellt. Die Rekonstruktion, mit einer Auflösung von 8,8Å (FSC=0,5), erlaubt das Einpassen von Homolologiemodellen und somit das Erstellen eines molekularen Modells mit pseudo atomarer Auflösung. Des Weiteren wurden zwei Rekonstruktionen des Hämocyanins von *Nautilus pompilius* (NpH) in verschiedenen Oxygenierungszuständen erstellt. Die auf 10 und 8,1Å aufgelösten Modelle zeigen zwei verschiedene Konformationen des Proteins. Daraus ließ sich eine Modellvorstellung über die allosterische Funktionsweise ableiten. Die hier erreichte Auflösung von 8Å ist die momentan höchste eines Molluskenhämocyanins.

Auf Grundlage des molekularen Modells von SoH konnte die Topologie des Proteins aufgeklärt werden. Es wurde gezeigt, dass die zusätzliche FU-d´ in den Kragen integriert ist und somit die prinzipielle Wandarchitektur aller Mollusken Hämocyanine identisch ist. Wie die Analyse des erstellten molekularen Modells zeigt werden sind die beiden Isoformen (SoH1 und SoH2) in den Bereichen der Interfaces nahezu identisch; auch der Vergleich mit NpH zeigt grosse Übereinstimmungen. Des weiteren konnte eine Fülle von Informationen bezüglich der allosterischen Signalübertragung innerhalb des Moleküls gewonnen werden.

Der Versuch, NpH in verschiedenen Oxygenierungszuständen zu zeigen, war erfolgreich. Die Datensätze, die unter zwei atmosphärischen Bedingungen präpariert wurden, führten reproduzierbar zu zwei unterschiedlichen Rekonstruktionen. Dies zeigt, daß der hier entwickelte experimentelle Ansatz funktioniert. Er kann nun routinemäßig auf andere Proteine angewandt werden. Wie der strukturelle Vergleich zeigte, verändert sich die Orientierung der FUs durch die Oxygenierung leicht. Dies wiederum beeinflusst die Anordnung innerhalb der Interfaces sowie die Abstände zwischen den beteiligten Aminosäuren. Aus dieser Analyse konnte eine Modellvorstellung zum allosterischen Signaltransfer innerhalb des Moleküls abgeleitet werden, die auf einer Umordnung von Salzbrücken basiert.

Literature

Al-Amoudi, A., Chang, J. J., Leforestier, A., McDowall, A., Salamin, L. M., Norlen, L. P., Richter, K., Blanc, N. S., Studer, D. and Dubochet, J. (2004a). "Cryo-electron microscopy of vitreous sections." EMBO J **23**(18): 3583-8.

Al-Amoudi, A., Norlen, L. P. and Dubochet, J. (2004b). "Cryo-electron microscopy of vitreous sections of native biological cells and tissues." J Struct Biol **148**(1): 131-5.

Boisset, N. and Mouche, F. (2000). "Sepia officinalis hemocyanin: A refined 3D structure from field emission gun cryoelectron microscopy." Journal of Molecular Biology(296 (2)): 459-472.

Bonaventura, C., Bonaventura, J., Miller, K. I. and Van Holde, K. E. (1981). "Hemocyanin of the chambered nautilus: structure-function relationships." Arch Biochem Biophys **211**(2): 589-98.

Bottcher, B., Kiselev, N. A., Stel'Mashchuk, V. Y., Perevozchikova, N. A., Borisov, A. V. and Crowther, R. A. (1997). "Three-dimensional structure of infectious bursal disease virus determined by electron cryomicroscopy." J Virol **71**(1): 325-30.

Bracewell, R.N. and Riddle, A.C. (1967). "Inversion of fan-beam scans into radio astronomy." Astrophys. Soc. **150**: 427-434Carragher, B., Kisseberth, N., Kriegman, D., Milligan, R. A., Potter, C. S., Pulokas, J. and Reilein, A. (2000). "Leginon: an automated system for acquisition of images from vitreous ice specimens." J Struct Biol **132**(1): 33-45.

Clare, D. K., Bakkes, P. J., van Heerikhuizen, H., van der Vies, S. M. and Saibil, H. R. (2009). "Chaperonin complex with a newly folded protein encapsulated in the folding chamber." Nature **457**(7225): 107-10.

Cong, Y., Zhang, Q., Woolford, D., Schweikardt, T., Khant, H., Dougherty, M., Ludtke, S. J., Chiu, W. and Decker, H. (2009). "Structural mechanism of SDS-induced enzyme activity of scorpion hemocyanin revealed by electron cryomicroscopy." Structure **17**(5): 749-58.

Connelli, P.R., Gill, S.J., Miller, K.I, Zhou, G. & and Van Holde, K. E. (1989). "Identical linkage and cooperativity of oxygen and carbon monoxide binding to *Octopus dofleini* hemocyanin " Biochemistry **28**: 1835-1843.

Cuff, M. E., Miller, K. I., van Holde, K. E. and Hendrickson, W. A. (1998). "Crystal structure

of a functional unit from Octopus hemocyanin." J Mol Biol **278**(4): 855-70.

Decker, H., Hellmann, N., Jaenicke, E., Lieb, B., Meissner, U. and Markl, J. (2007). "Minireview: Recent progress in hemocyanin research." Integrative and Comparative Biology **47**(4): 631-644.

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

Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, A. W. and Schultz, P. (1988). "Cryo-electron microscopy of vitrified specimens." Q Rev Biophys **21**(2): 129-228.

Dubochet, J., Lepault, J., Freeman, R., Berriman, J.A. and and Homo, J. C. (1982). "Electron microscopy of frozen water and aqueous solutions." J. Microsc. **128**: 219-237.

Elad, N., Clare, D. K., Saibil, H. R. and Orlova, E. V. (2008). "Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections." J Struct Biol **162**(1): 108-20.

Fernandez, J. J., Luque, D., Caston, J. R. and Carrascosa, J. L. (2008). "Sharpening high resolution information in single particle electron cryomicroscopy." J Struct Biol **164**(1): 170-5.

Fiser, A. and Sali, A. (2003). "Modeller: generation and refinement of homology-based protein structure models." Methods Enzymol **374**: 461-91.

Frank, Joachim (2006). "Three-Dimensional Electron Microscopy of Macromolecular Assemblies " Oxford University Press.

Gatsogiannis, C., Moeller, A., Depoix, F., Meissner, U. and Markl, J. (2007). "Nautilus pompilius hemocyanin: 9 angstrom cryo-TEM structure and molecular model reveal the subunit pathway and the interfaces between the 70 functional units." Journal of Molecular Biology **374**(2): 465-486.

Gatsogiannis, Christos and Markl, Jürgen (2008). "Keyhole Limpet Hemocyanin: The 9-Å CryoEM Structure and Molecular Model of the KLH1 Didecamer Reveal the Interfaces and Intricate Topology of 160 Functional Units." Journal of Molecular Biology **385**(3): 963-83.

Gielens, C., De Geest, N., Xin, X. Q., Devreese, B., Van Beeumen, J. and Preaux, G. (1997). "Evidence for a cysteine-histidine thioether bridge in functional units of molluscan haemocyanins and location of the disulfide bridges in functional units d and g of the betaC-haemocyanin of Helix pomatia." Eur J Biochem **248**(3): 879-88.

Glaeser, Robert M. and Taylor, K.A. (1978). "Radiation damage relative to transmisson electron microscopy of biological specimens at low temperature: a review." J. Microsc. **122**: 127-138

Harauz, G. and Van Heel, M. (1986). “Exact filters for general geometry three-dimensional reconstruction.” Optik **73**: 146-156.

Harris, J. R., Gebauer, W., Guderian, F. U. and Markl, J. (1997). “Keyhole limpet hemocyanin (KLH), I: Reassociation from Immucothel followed by separation of KLH1 and KLH2.” Micron **28**(1): 31-41.

Harris, J. R., Meissner, U., Gebauer, W. and Markl, J. (2004). “3D reconstruction of the hemocyanin subunit dimer from the chiton Acanthochiton fascicularis.” Micron **35**(1-2): 23-26.

Hartmann, H., Bongers, A. and Decker, H. (2003). “Small-angle X-ray Scattering-based Three-dimensional Reconstruction of the Immunogen KLH1 Reveals Different Oxygendependent Conformations.” Journal of Biological Chemistry **279**(4): 2841-2845.

Hong, Z. and Zhou (2008). “Towards atomic resolution structural determination by single-particle cryo-electron microscopy “ Curr Opin Struct Biol **18**(2): 126-134.

Klaholz, B. P., Myasnikov, A. G. and Van Heel, M. (2004). “Visualization of release factor 3 on the ribosome during termination of protein synthesis.” Nature **427**(6977): 862-5.

Lambert, O., Boisset, N., Taveau, J-C. and Lamy, J. N. (1994). “Three-Dimensional Reconstruction of Sepia officinalis Hemocyanin from Frozen-Hydrated Specimens.” Archives of Biochemistry and Biophysics **316**(No.2): 950-959.

Lang, W. H. (1988). “cDNA cloning of the Octopus dofleini hemocyanin: sequence of the carboxyl-terminal domain.” Biochemistry **27**(19): 7276-82.

Lang, W. H. and van Holde, K. E. (1991). “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.

Llorca, O., Smyth, M. G., Carrascosa, J. L., Willison, K. R., Radermacher, M., Steinbacher, S. and Valpuesta, J. M. (1999). “3D reconstruction of the ATP-bound form of CCT reveals the asymmetric folding conformation of a type II chaperonin.” Nat Struct Biol **6**(7): 639-42.

Ludtke, S. J., Baker, M. L., Chen, D. H., Song, J. L., Chuang, D. T. and Chiu, W. (2008). “De novo backbone trace of GroEL from single particle electron cryomicroscopy.” Structure **16**(3): 441-8.

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

Markl, J. and Decker, H. (1992). “Molecular structure of the arthropod hemocyanins.”

Adv In Comp & Envir Physiol **13**(325-376).

Markl, J., Moeller, A., Martin, A. G., Rheinbay, J., Gebauer, W. and Depoix, F. (2009). "10-A CryoEM Structure and Molecular Model of the Myriapod (Scutigera) 6x6mer Hemocyanin: Understanding a Giant Oxygen Transport Protein." J Mol Biol.

Martin, A. G., Depoix, F., Stohr, M., Meissner, U., Hagner-Holler, S., Hammouti, K., Burmester, T., Heyd, J., Wriggers, W. and Markl, J. (2007). "Limulus polyphemus hemocyanin: 10 angstrom cryo-TEM structure, sequence analysis, molecular modelling and rigid-body fitting reveal the interfaces between the eight hexamers." Journal of Molecular Biology **366**(4): 1332-1350.

Meissner, U., Dube, P., Harris, J. R., Stark, H. and Markl, J. (2000). "Structure of a molluscan hemocyanin didecamer (HtH1 from Haliotis tuberculata) at 12 angstrom resolution by cryoelectron microscopy." Journal of Molecular Biology **298**(1): 21-34.

Meissner, U., Gatsogiannis, C., Moeller, A., Depoix, F., Harris, J. R. and Markl, J. (2007). "Comparative 11 angstrom structure of two molluscan hemocyanins from D cryo-electron microscopy." Micron **38**(7): 754-765.

Mellema, J. E. and Klug, A. (1972). "Quaternary structure of gastropod haemocyanin." Nature **239**(5368): 146-50.

Miller, K. I., Cuff, M. E., Lang, W. F., Varga-Weisz, P., Field, K. G. and van Holde, K. E. (1998). "Sequence of the Octopus dofleini hemocyanin subunit: structural and evolutionary implications." J Mol Biol **278**(4): 827-42.

Miller, K. I., van Holde, K. E., Toumadje, A., Johnson, W. C., Jr. and Lamy, J. (1988). "Structure and function of the carboxyl-terminal oxygen-binding domain from the subunit of Octopus dofleini hemocyanin." Biochemistry **27**(19): 7282-8.

Miller, Karen I. (1984). "Oxygen Equilibria of *Octopus dofleini* Hemocyanin." Biochemistry **24**: 4586-4592.

Mindell, J. A. and Grigorieff, N. (2003). "Accurate determination of local defocus and specimen tilt in electron microscopy." J Struct Biol **142**(3): 334-47.

Monod, J., Wyman, J. and Changeux, J-P. (1965). "On the Nature of Allosteric Transitions: A Plausible Model." JMB **12**(1): 88-118.

Orlova, E. V., Dube, P., Harris, J. R., Beckman, E., Zemlin, F., Markl, J. and van Heel, M. (1997a). "Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 A resolution by electron cryomicroscopy and angular reconstitution." J Mol Biol **271**(3): 417-37.

Orlova, E. V., Dube, P., Harris, J. R., Beckman, E., Zemlin, F., Markl, J. and vanHeel, M. (1997b). "Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 angstrom resolution by electron cryomicroscopy and angular reconstitution." Journal of

Molecular Biology **271**(3): 417-437.

Penczek, P. A., Frank, J. and Spahn, C. M. (2006). "A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation." J Struct Biol **154**(2): 184-94.

Penzcek, P.A., Grassucci, R.A. and Frank, J. (1994). "The ribosome at improved resolution: New techniques for mergin and orientation refinement in 3D cryoelectron microscopy " Ultramicroscopy **53**: 251-270.

Perbrandt, M., Guthohrlein, E.W., Rypniewski, W., Idakieva, K., Stoeva, S., Voelter, W., Genov, N. & and Betzel, C. (2003). "The structure of a functional unit from the wall of a gastropod hemocyanin offers a possible mechanism for cooperativity." Biochemistry **42**: 6341-6346.

Potter, C. S., Chu, H., Frey, B., Green, C., Kisseberth, N., Madden, T. J., Miller, K. L., Nahrstedt, K., Pulokas, J., Reilein, A., Tcheng, D., Weber, D. and Carragher, B. (1999). "Leginon: a system for fully automated acquisition of 1000 electron micrographs a day." Ultramicroscopy **77**(3-4): 153-61.

Radermacher, M. (1988). "Three-dimensional reconstruction of single particles from random and nonrandom tilt series." J Electron Microsc Tech **9**(4): 359-94.

Radermacher, M., Rao, V., Grassucci, R., Frank, J., Timerman, A. P., Fleischer, S. and Wagenknecht, T. (1994). "Cryo-electron microscopy and three-dimensional reconstruction of the calcium release channel/ryanodine receptor from skeletal muscle." J Cell Biol **127**(2): 411-23.

Radon, J. (1917). "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig." Math. Phys. Klasse **69**: 262-277.

Ranson, N. A., Farr, G. W., Roseman, A. M., Gowen, B., Fenton, W. A., Horwich, A. L. and Saibil, H. R. (2001). "ATP-bound states of GroEL captured by cryo-electron microscopy." Cell **107**(7): 869-79.

Roseman, A. M., Ranson, N. A., Gowen, B., Fuller, S. D. and Saibil, H. R. (2001). "Structures of unliganded and ATP-bound states of the Escherichia coli chaperonin GroEL by cryoelectron microscopy." J Struct Biol **135**(2): 115-25.

Schatz, M., Orlova, E. V., Dube, P., Jager, J. and van Heel, M. (1995). "Structure of Lumbricus terrestris hemoglobin at 30 A resolution determined using angular reconstitution." J Struct Biol **114**(1): 28-40.

Scheres, S. H. and Carazo, J. M. (2009). "Introducing robustness to maximum-likelihood refinement of electron-microscopy data." Acta Crystallogr D Biol Crystallogr **65**(Pt 7): 672-8.

Scheres, S. H., Valle, M., Grob, P., Nogales, E. and Carazo, J. M. (2009). "Maximum likelihood refinement of electron microscopy data with normalization errors." J Struct Biol **166**(2): 234-40.

Senozan, N. M., Landrum J, Bonaventura, J. and Bonaventura, C. (1981). "Hemocyanin of the giant keyhole limpet *Megathura crenulata*." Dekker **Oxygen Binding Proteins** 703-717.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C. S. and Carragher, B. (2005). "Automated molecular microscopy: the new Leginon system." J Struct Biol **151**(1): 41-60.

Van Heel, M. (1987). "Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction." Ultramicroscopy **21**: 111-124.

Van Heel, M. and Frank, J. (1981). "Use of multivariate statistical statistics in analysing the images of biological macromolecules." Ultramicroscopy **9**: 187-194.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M. and Patwardhan, A. (2000). "Single-particle electron cryo-microscopy: towards atomic resolution." Q Rev Biophys **33**(4): 307-69.

van Heel, M., Harauz, G. and Orlova, E. V. (1996). "A new generation of the IMAGIC image processing system." J Struct Biol **116**(1): 17-24.

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

Van Holde, K. E. (1967). "Physical studies of hemocyanins. 3. Circular dichroism and absorption spectra." Biochemistry **24**(17): 4577-4582.

van Holde, K. E., Miller, K. I. and Decker, H. (2001). "Hemocyanins and invertebrate evolution." J Biol Chem **276**(19): 15563-6.

van Holde, K. E., Miller, Karen I., C.B. Anfinsen, Frederic M. Richards John T. Edsall and David, S. Eisenberg (1995). Hemocyanins. Advances in Protein Chemistry, Academic Press. **Volume 47:** 1-81.

Ward, J.H., Jr. (1982). "Hierachical grouping to optimize an objective function." Am. Statist. Assoc. J. **58**: 236-244.

Westermann, B., Beck-Schildwachter, I., Beuerlein, K., Kaleta, E. F. and Schipp, R. (2004). "Shell growth and chamber formation of aquarium-reared Nautilus pompilius (Mollusca, Cephalopoda) by X-ray analysis." J Exp Zoolog A Comp Exp Biol **301**(12): 930-7.

Zhang, W., Kimmel, M., Spahn, C. M. and Penczek, P. A. (2008). "Heterogeneity of large macromolecular complexes revealed by 3D cryo-TEM variance analysis." Structure **16**(12): 1770-6.

Zolla, L., Brunori, M., Richey, B. and Gill, S. J. (1985). "Heterogeneous binding of oxygen

and carbon monoxide to dissociated molluscan hemocyanin." Biophys Chem **22**(4): 271-80.

Zolla, L., Kuiper, H. A., Vecchini, P., Antonini, E. and Brunori, M. (1978). "Dissociation and oxygen-binding behaviour of beta-hemocyanin from Helix pomatia." Eur J Biochem **87**(3): 467-73.

List of abbrevations

2D	two-dimensional
3D	three-dimensional
Å	Ångstrom (10^{-10} meter)
C5	fivefold rotational symmetry
D5	dihedral fivefold symmetry
Da	Dalton
FSC	Fourier Shell Correlation
FT	Fourier Transformation
FU	Functional Unit
m-r-a	multi-reference-alignment
MSA	Multivariate Statistical Analysis
NpH	*Nautilus pompilius* Hemocyanin
PDB	Protein Data Bank
pH	potentia Hydrogenii
SAXS	Small Angle X-ray Scattering
SNR	Signal to Noise Ratio
SoH	*Sepia officinalis* Hemocyanin
tif(f)	tagged image (file) format

Amino acids:

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